Performance of GLCM Algorithm for Extracting Features to Differentiate Normal and Abnormal Brain Images

Zul Indra¹, Yessi Jusman²,*

¹Department of Informatics Engineering, Faculty of Engineering, Abdurrab University, Kota Pekanbaru, Riau 28291, Indonesia
²,*Department of Electrical Engineering, Faculty of Engineering, Universitas Muhammadiyah Yogyakarta, Bantul, Daerah Istimewa Yogyakarta 55183, Indonesia

Corresponding author E-mail: yjusman@umy.ac.id

Abstract. Brain cancer is a malignant brain tumor that can spread quickly to other parts of the brain and spine. However, not all tumors are malignant and can be treated before they become malignant. The purpose of this study is to discover brain abnormalities based on CT scan images by using T-test algorithm. Thus, it can be one of solution for early detection of brain abnormalities in order to treat it before it becomes a malignant tumor or cancer. As dataset, this research using 40 images consisting of 20 normal brain images and 20 abnormal brain images. There are two algorithms which are used in this research i.e. Gray level co-occurrences matrix (GLCM) for feature extraction and T-Test for brain image classification. Prior to feature extraction, brain image is converted to Graycomatrix in order to adjust the brightness of the image. The final step is image classification by using the T-test algorithm. From 40 test results which are used in this study, GLCM method can extract 8 features that can significantly distinguish the image of normal brain and abnormal brain. For the T-test algorithm, it is found that each feature has a P-value <0.05 which means that extracted features can be used for the further classification process of brain image abnormality. Thus, it can be inferred that this research framework which is employed the GLCM and T-test algorithm can be used to assist the process of early diagnosis of brain cancer.

1. Introduction

Brain tumors is a of main public health problem which is one of the biggest causes of death in the world [1]. Due to level of mortality, brain tumour is the second dangerous disease that causes death for men aged 20-30 years and is the fifth dangerous disease that causes death for women aged 20-30 years. In 2010, International Agency for Research on Cancer declared that more than 126,000 people in the world each year have brain tumour and more than 97,000 people die [2]. In addition to having a very high mortality rate, brain cancer is also a very dangerous disease in terms of the difficulty to make an initial diagnosis [3]. In the early stages, the brain tumour is very difficult to be discovered because the tumour boundary is still unclear, the contrast is too low and sometimes resembles as a normal tissue [4].

Brain cancer is a malignant brain tumour that can spread rapidly to the other parts of the brain and spine [5]. However, not all brain tumours are malignant and can be categorized as cancer [6]. There are also benign brain tumours. Benign brain tumours are a group of brain cells that grow slowly and do not spread to the other parts [7]. The brain tumours itself is an abnormal and uncontrolled
growth of brain cells. In the brain, tumour can develop from cells that make up brain tissue, from nerves that come in and out of the brain, and from the protective membranes of the brain and spinal cord. Brain cancer is an event where abnormal tissue growth occurs in the brain [5]. Abnormal tissue growth in the brain can be indicated by symptoms that appear, including vomiting, seizures, decreased nerve function, and emotional. In order to treat brain cancer effectively, there is a need for the right strategy to deal with it. One of strategy that can be done is to help the process of early diagnosis of brain tumours [8, 9].

It can be inferred that early diagnosis is very necessary in terms of handling brain tumour to prevent further damage and even death for the patient. Along with the rapid development of technology, one technique that is quite widely applied to help diagnose diseases is the digital image processing techniques. Digital image processing is utilized to analyse the medical images which are produced by CT scans or magnetic imaging resonance (MRI). However, those medical images are very high-resolution image so that it requires a lot of time to be examined by experts. To overcome this time-consuming problem, a feature selection method was presented in the digital image processing technique. Feature selection is a pre-processing step to reduce the complexity of the image during diagnostic models [9]. The main goal of the feature extraction is to discover the features which is the best representation and contain fewer parameter for medical images [10, 11]. By finding the right features, the process of diagnosing brain tumours can be performed faster because non-essential features in medical images of the brain have been successfully eliminated.

This research is intended to propose an intelligence system which can assist the early diagnosis of brain tumour detection by providing training data for further classification system. These training data are obtained from the feature extraction process of medical image by using GLCM algorithm. Gray level co-occurrences matrix (GLCM) is one most popular texture-based feature extraction algorithm which can assign the textural relationship between the pixels images. The main principle of the GLCM is to perform operations according to second-order statistics in the image [12, 13, 14]. The research framework for this study consists of two main stages namely features extraction of the brain image and the validation of these extracted features. A set of feature matrix which can represent the medical image with fewer parameters is the final output from this features extraction stage. After that, these feature matrixes are then validated by using T-Test algorithm and P-Value standard.

2. Related Work

The main challenge of feature extraction is to determine the appropriate features in order to reduce the complexity of medical images without reducing the quality of the information in medical images. Up to this time, there have been several studies that have been attempted to be carried out regarding features extraction of brain images. The summary of the related work for feature extraction of brain tumour is explained in Table 1.

From Table 1, it can be inferred that the research to find the best way in extracting brain image feature is still ongoing today. Furthermore, until now, there has not been one study that conducted research to apply data validation of the generated features. Therefore, this research proposed a new approach for extracting the brain images feature. This effort is hoped will increase the quality the feature extraction process as well as to assist the early diagnosis of brain tumour detection.

Table 1. Work Related to Feature Extraction of Brain Tumour Image

| Author Name and Year | Feature Extraction Algorithm | Feature Reduction and Optimization | Feature Validation |
|----------------------|-------------------------------|----------------------------------|--------------------|
| Sharma et al. (2014) [15] | GLCM                          | No                               | No                 |
| Sudha et al. (2014) [16] | GLCM and GLRM                 | No                               | No                 |
| Demirhan et al (2014) [17] | SWT                           | No                               | No                 |
| Gaikwad, S. B., & Joshi, M (2015) [18] | PCA                           | No                               | No                 |
| Vidyarthi, A., & Mittal, N (2015) | Gabor wavelet and DWT         | No                               | No                 |
3. Methodology

3.1. Research Framework
The research framework for this study is divided into two main stages namely features extraction of the brain image and validation of the extracted features. The flowchart for this research is illustrated in Figure 1.

![Figure 1. Research Framework.](image)

3.2. Data Collection
As in general, data collection is the first section for this research. In this research, data collection is performed by collecting the brain images from Radiopedia, a wiki-based international collaborative radiology educational web resource. To conduct an experiment, this study uses two types of images, namely normal and abnormal brain images. The total dataset of this study is 20 brain images for each type of image. In addition, these used brain images is captured in 904 x 1024 pixels. Furthermore, each of images is then captured in ten times for different positions.

3.3. Feature Extraction
The next process after data collection is feature extraction as presented in Figure 2. Feature extraction is the process of taking the characteristics contained in an object of the image. However prior to the feature extraction section, the sampling image of the brain is converted into grayscale image to adjust the brightness of the image. Grey Level Co-Concurrent Matrix (GLCM) algorithm is employed to carry out this feature extraction. Feature extraction with texture analysis is aimed to obtain the
characteristics of a grayscale image in order to distinguish one image from another image. In this study, the image is extracted based on four characteristics namely entropy parameters, contrast, energy, correlation. Each characteristic is an important feature because it contains valuable information in an image. These characteristics can be considered as the surface structure of an image.

![Feature Extraction Flowchart](image)

**Figure 2.** Flowchart of Feature Extraction.

3.4. Feature Extraction

The last section in this study is validating the extracted features as presented in Figure 3. This section is intended to validate the feature whether the feature can be used for the further classification process of brain image abnormality. This study utilizes the T-test algorithm to achieve this purpose [27,28,29,30]. Apart from that, the P-Value standard is used to carry out this feature validation process [31,32].
4. Results & Analysis

4.1. Result of Feature Extraction

As explained before, there are two types of images which are used in this study i.e. normal and abnormal brain images. Both Normal and abnormal images have differences texture that can be seen by naked eye. This difference becomes a reference to determine the type of normal and abnormal images. The difference in terms of texture can be seen in the Figures 4 and 5.

Prior to extract the feature, the brain images which are used as the dataset have to be converted into Graycomatrix form. This conversion process is intended so that the image has a better contrast value. Comparison of brain images in Grayscale form with brain images that have been changed to Graycomatrix. The value will make it easier for us to distinguish between normal brain cells and abnormal brain cells. Once the conversion process of the image into Graycomatrix is completed, the next step is to extract feature of the converted Graycomatrix images. As explained before, the feature of image is extracted based on four entropy characteristics namely contrast, energy, correlation and homogeneity.

Feature extraction can be varied according to the displacement and orientation value [33,34,35,36,37]. There are two displacement values that can be used namely \( d = 1 \) and \( d = 2 \) and four orientation values \( \theta = 0^\circ, 45^\circ, 90^\circ, \) and \( 135^\circ \) to form the GLCM matrix. The results of the feature extraction and its variations can be seen in Table 2.
Table 2. Extracted Feature

| Angle Used | Name of Feature             |
|------------|-----------------------------|
| 4 Angles   | Contrast1, Corelation1, Energy1, Homogeneity1 |
| 8 Angles   | Contrast1, Contrast2, Corelation1, Corelation2, Energy1, Energy2, Homogeneity1, Homogeneity2 |

From Table 2, it can be seen that there are several features produced from the feature extraction process according to the used of displacement values. When only $d=1$ displacement value is used, it will produce four features namely contrast1, corellation1, energy1, homogeneity1. However, if both of the two of displacement value ($\theta d=1$ and $d=2$) are used, the GLCM algorithm will produced eight feature. The displacement value $d=1$ produces four features namely contrast1, correlation1, energy1 and homogeneity1. Four other features namely contrast2, correlation2, energy2 and homogeneity2 are obtained from the use of the displacement value $d=2$. Once the value of each feature characteristic is obtained, the feature extraction process is completely done. The next section is the validation of these extracted values.

4.2. Result of Feature Validation

The validation of the extracted features is the last section in this study which is intended to validate the feature whether the feature can be used for the classification process of brain image abnormality. The validation process is performed by using T-test algorithm (i.e. mean, standard deviation, and P values). The result of this validation process is described in the following Tables 3 and 4. From Tables 3 and 4, it is discovered that all features extracted using GLCM Algorithm have P-value lower than 0.05 where the value of 0.05 is the threshold of the features feasibility. Thus it can be concluded that all of the extracted features are proper for the brain cell classification.

Table 3. Feature Validation for 4 Angles

| Feature | Image Type | Mean | St.Dev | P-Value |
|---------|------------|------|--------|---------|
| Contrast1 | Abnormal | 0,07 | 0,02 | 0,000 |
| Normal | 0,11 | 0,02 |
| Abnormal | 0,99 | 0,00 |
| Correlation1 | Normal | 0,99 | 0,00 | 0,000 |
| Abnormal | 0,34 | 0,03 |
| Energy1 | Normal | 0,26 | 0,04 | 0,000 |
| Homogeneity1 | Abnormal | 0,96 | 0,00 | 0,000 |

Table 4. Feature Validation for 8 Angles

| Feature | Image Type | Mean | St.Dev | P-Value |
|---------|------------|------|--------|---------|
| Contrast1 | Abnormal | 0,65 | 0,02 | 0,000 |
| Normal | 0,10 | 0,02 |
| Abnormal | 0,10 | 0,03 |
| Normal | 0,17 | 0,04 |
| Contrast2 | Abnormal | 0,99 | 0,00 | 0,000 |
| Normal | 0,99 | 0,00 |
| Correlation1 | Abnormal | 0,99 | 0,00 | 0,000 |
| Normal | 0,99 | 0,00 |
| Correlation2 | Abnormal | 0,34 | 0,03 | 0,000 |
| Normal | 0,26 | 0,04 | 0,000 |
Energy2 | Abnormal | 0.34 | 0.04 | 0.000
| Normal  | 0.26 | 0.06 | 0.000

Homogeneity1 | Abnormal | 0.98 | 0.00 | 0.007
| Normal  | 0.97 | 0.00 | 0.007

Homogeneity2 | Abnormal | 0.97 | 0.05 | 0.004
| Normal  | 0.96 | 0.01 | 0.004

5. Conclusion
Based on the results of the research that has been conducted, it proved that the Gray level co-occurrences matrix (GLCM) method can be used to extract features to distinguish normal and abnormal brain images. The features extracted in this study are based of four characteristics i.e. contrast, correlation, energy and homogeneity. All features extracted using GLCM Algorithm have P-value lower than 0.0 where the value of 0.05. It can be inferred that this research framework which is employed the Gray level co-occurrences matrix (GLCM) and T-test algorithm can be used to assist the process of early diagnosis.

References
[1] Dora L, Agrawal S, Panda R, and Abraham A, 2018, Nested cross-validation based adaptive sparse representation algorithm and its application to pathological brain classification, Expert Syst. Appl., 114, 313–321.
[2] Al-Tamimi M S H, and Sulong G, 2014, Tumor Brain Detection Through Mr Images: A Review Of Literature, J. Theor. Appl. Inf. Technol., 62(2).
[3] Özüyurt F, Sert E, Avci E, and Dogantekin E, 2019, Brain tumor detection based on Convolutional Neural Network with neutrosophic expert maximum fuzzy sure entropy, Measurement, 147, 106830.
[4] Tong J, Zhao Y, Zhang P, Chen L, and Jiang L, 2019, MRI brain tumor segmentation based on texture features and kernel sparse coding, Biomed. Signal Process. Control, 47, 387–392.
[5] Shanthakumar P, and Ganeshkumar P, 2015, Performance analysis of classifier for brain tumor detection and diagnosis, Comput. Electr. Eng., 45, 302–311.
[6] Sachdeva J, Kumar V, Gupta I, Khandelwal N, and Ahuja C K, 2016, A package-SFERCB-“Segmentation, feature extraction, reduction and classification analysis by both SVM and ANN for brain tumors, Appl. Soft Comput. J., 47, 151–167.
[7] Abd-Ellah M K, Awad A I, Khalaf A A M, and Hamed H F A, 2019, A review on brain tumor diagnosis from MRI images: Practical implications, key achievements, and lessons learned, Magn. Reson. Imaging, 61, 300–318.
[8] Mohan G, and Subashini M M, 2018, MRI based medical image analysis: Survey on brain tumor grade classification, Biomed. Signal Process. Control, 39, 139–161.
[9] Kumar A, Ramachandran M, Gandomi A H, Patan R, Lukasis S, and Soundarapandian R K, 2019, A deep neural network based classifier for brain tumor diagnosis, Appl. Soft Comput. J., 82, 105528.
[10] Öztürk Ş, and Akdemir B, 2018, Application of Feature Extraction and Classification Methods for Histopathological Image using GLCM, LBP, LBGLCM, GLRLM and SFTA, Procedia Comput. Sci., 132, 40–46.
[11] Remeseiro B, and Bolon-Canedo V, 2019, A review of feature selection methods in medical applications, Comput. Biol. Med., 112, 103375.
[12] Khalil M, Ayad H, and Adib A, 2018, Performance evaluation of feature extraction techniques in MR-Brain image classification system, Procedia Comput. Sci., 127, 218–225.
[13] Arabi P M, Joshi G, and Vamsha Deepa N, 2016, Performance evaluation of GLCM and pixel intensity matrix for skin texture analysis, Perspect. Sci., 8, 203–206.
[14] Abd Latif M H, Md Yusof H, Sidek S N, and Rusli N, 2015, Implementation of GLCM Features in Thermal Imaging for Human Affective State Detection, Procedia Comput. Sci., 76, 308–315.
[15] Sharma K, Kaur A, and Gujral S, 2014, Brain tumor detection based on machine learning
algorithms, Int. J. Comput. Appl., 103 (1).
[16] Sudha B, Gopikannan P, Shenbagarajan A, and Balasubramanian C, 2014, Classification of Brain Tumor Grades using Neural Network, in Proceedings of the World Congress on Engineering 2014, 1.
[17] Demirhan A, Törü M, and Güler I, 2014, Segmentation of tumor and edema along with healthy tissues of brain using wavelets and neural networks, IEEE J. Biomed. Heal. informatics, 19(4), 1451–1458.
[18] Gaikwad S B, and Joshi M S, 2015, Brain tumor classification using principal component analysis and probabilistic neural network, Int. J. Comput. Appl., 120(3).
[19] Vidyarthi A, and Mittal N, 2015, Performance analysis of Gabor-Wavelet based features in classification of high grade malignant brain tumors, in 2015 39th National Systems Conference (NSC), 1–6.
[20] Kumar S P M, and Chattejee S, 2017, Computer aided diagnostic for cancer detection using MRI images of brain (Brain tumor detection and classification system), 2016 IEEE Annu. India Conf. INDICON 2016.
[21] Banday S A, and Mir A H, 2016, Statistical textural feature and deformable model based MR brain tumor segmentation, 2016 Int. Conf. Adv. Compt. Commun. Informatics, ICACCI 2016, 657–663.
[22] R M A, and Scholar P G, 2017, Brain Tumor Segmentation And Classification using DWT, Gabour Wavelet And GLCM, 1744–1750.
[23] Islam A, Hossain M F, and Saha C, 2018, A new hybrid approach for brain tumor classification using DWT-KSVM, 4th Int. Conf. Adv. Electr. Eng. ICAEE 2017, 241–246.
[24] Sergaki E. et al., 2018, Application of ANN and ANFIS for detection of brain tumors in MRIs by using DWT and GLCM texture analysis, IST 2018 - IEEE Int. Conf. Imaging Syst. Tech. Proc., 1–6.
[25] Dubey Y K, Mushrif M M, and Pisar K, 2018, Brain tumor type detection using texture features in MR images, IEEE Reg. 10 Humanit. Technol. Conf. R10-HTC, 1–4.
[26] Gopika G S, and Rajasree R S, 2018, Certain Investigations on the Detection of Brain Tumors Using Support Vector Machine & Prinicipal Component Analysis,” in 2018 Fourth International Conference on Computing Communication Control and Automation (ICCCUBEA), 1–5.
[27] Kim T K, 2019, T-test as a parametric statistic, Korean J. Anesthesiol., 68(6), 540.
[28] Williams L L, and Quave K, 2019, Comparing Two Groups: t-Tests, Quant. Anthropol., 89–104.
[29] Feng Y, Huang Y, and Ma X, 2017, The application of Student’s t-test in internal quality control of clinical laboratory, Front. Lab. Med., 1(3), 125–128.
[30] Smallheiser N R, 2017, Null Hypothesis Statistical Testing and the t-Test, Data Lit., 127–136.
[31] Buas M F, Li C I, Anderson G L, and Pepe M S, 2018, Recommendation to use exact P-values in biomarker discovery research in place of approximate P-values, Cancer Epidemiol., 56, 83–89.
[32] Bergamelli M, Bianchi A, Khalaf L, and Urga G, 2019, Combining p-values to test for multiple structural breaks in cointegrated regressions, J. Econom., 211(2), 461–482.
[33] Jusman Y, Ng S C, Hasikin K, Kurnia R, Abu Osman N A, and Teoh K H, 2017, A system for detection of cervical precancerous in field emission scanning electron microscope images using texture features, J. Innov. Opt. Health Sci., 10 (2), 1650045.
[34] Jusman Y, Ng S C, Hasikin K, Kurnia R, Abu Osman N A, and Teoh K H, 2016, Computer-aided screening system for cervical precancerous cells based on field emission scanning electron microscopy and energy dispersive x-ray images and spectra, Opt. Eng., 55(10), 103110.
[35] Jusman Y, S.-C. Ng, and K. Hasikin “Performances of Proposed Normalization Algorithm for Iris Recognition,” Int. J. Adv. Intell. Informatics, Vol. 6, no. 2, July 2020.
[36] Jusman, L. A. Dewiprabamukti, A. N. N. Chaimim, Z. Mohamed, S. N. A. M. Kanaﬁah, & N. H. A. Halim. "Application of Watershed Algorithm and Gray Level Co-Occurrence Matrix in Leukemia Cells Images”. IEEE. Paper presented at the 2020 3rd MECnIT. pp. 9-14, 2020.
[37] Jusman Y, Tamarena R I, Puspita S, Saleh E, and Kanaﬁah S N A M, 2020, Analysis of Features Extraction Performance to Differentiate of Dental Caries Types Using Gray Level Co-occurrence Matrix Algorithm, Paper presented at the 2020 10th IEEE ICSCSC, 148-152.