ISSN 2745-3502
Volume 3 No. 2 – December 2022 *Indonesian Applied Physics Letters*

# Identification of Stroke with MRI Images Using the Learning Vector Quantization (LVQ) Method Based on Texture Features

Lellen Novia Hariono[1], Dr. Suryani Dyah Astuti[1], M.Si, Endah Purwanti, S.Si., M.T[1]

[1]Department of Physics, Faculty of Science and Technology, Universitas Airlangga Surabaya, East Java, Indonesia

[a)] email: liapriliyah@gmail.com

**Abstract.** A fire location monitoring system is designed in this study to determine the location of a fire on a ship. The inputs used are temperature detectors and smoke detectors. Research on the Identification of Stroke with MRI Imagery Using the Learning Vector Quantization (LVQ) Method Based on Texture Features has been carried out. This study aims to determine the program's best parameters and the highest accuracy level of the stroke identification program. This research was conducted at Haji Sukolilo Hospital - Surabaya by obtaining 57 images of stroke patients and 15 images of regular patients. The study used the intelligence of stroke, tumor, and standard images to determine each category's image characteristics. After knowing the differences in each class, the next process is digital image processing, followed by feature extraction used is the Gray-Level Co-occurrence matrix (GLCM) with four parameters: contrast, correlation, diversity, and homogeneity. These four parameters are the best input parameters with an intelligence rate of 0.100 with a decrease in intelligence rate of 0.100, so the best accuracy value for training is 74.97%, and test data is 78.60%. Regarding the program's ability to correctly identify 11 data from 14 data tested, the program is feasible to be used as a second opinion.

## INTRODUCTION

The brain consists of millions of nerve cells arranged in specific patterns that control emotions, behavior, body movements, and sensations. The development of the brain allows for damage. Brain damage occurs due to several things, such as traumatic brain injury (Traumatic Brain Injury), chemical damage to the brain, and acquired brain injury (Acquired Brain Injury). An example of a disease that causes brain damage is a stroke (cerebrovascular accident). A stroke (cerebrovascular accident) can cause a rupture of brain vessels so that the bleeding that occurs disrupts brain function. Stroke can be diagnosed with images produced by diagnostic equipment such as a CT scan (Computerized Tomography) or MRI (Magnetic Resonance Imaging). To avoid subjective nature, Computer Vision or computer vision can combine image processing with artificial intelligence. Computer Vision is a method that interprets an image to obtain the desired information accurately. One of the methods found in artificial neural networks is Learning Vector Quantization (LVQ). LVQ input is received from feature extraction of the Gray Level Co-occurrence Matrix.

Magnetic Resonance Imaging (MRI) is a medical tool in the field of radiological diagnostic examinations, which produces images of cross-sectional images of the human body/organs using a magnetic field of strength between 0.064 – 1.500 tesla (1 tesla = 1000 Gauss) and vibrational resonance to the hydrogen atomic nucleus.

The main advantage is that it can make coronal, sagittal, axial, and oblique cuts without much manipulating the patient's body position, so it is very suitable for soft tissue diagnostics (Notosiswoyo in Tjahjadi, 2014)

The brain is a highly specialized organ. The total weight of the adult brain is 2,000% of the entire body weight, or 1,400 kg. Information processing in the brain is carried out in particular parts according to the translation area of sensory neurons. The shaped surface of the brain is not flat but squiggly as the neurons develop in it.

Stroke is acutely a focal and general functional brain disorder disease (Riyadina et al., 2013). Stroke is a disease due to cerebral circulation disorders that are influenced by many risk factors consisting of those that

cannot be changed in the form of age and sex and those that can be changed, such as hypertension, increased blood sugar levels, dyslipidemia, and occupation. (Chintya 2013 in Riyadina et al, 2013).

An ischemic condition or hemorrhagic process causes a stroke, often preceded by a lesion or injury to the arteries. Of all stroke events, two-thirds are ischemic, and one-third are hemorrhagic. It is called an ischemic stroke because there is a blockage of blood vessels by thromboembolic, which results in the area under the backup experiencing ischemic. This differs from hemorrhagic strokes, which occur due to a ruptured microaneurysm (Riyadina et al., 2013). With the help of MRI, some damage to brain tissue can be identified through the results of MRI images.

Image processing refers to dimensional processing using a computer. In a broader context, digital image processing refers to processing any 2-dimensional data.

Texture analysis is an intermediate process for classifying and interpreting images. An image classification process based on texture analysis generally requires feature extraction stages consisting of statistical, spectral, and structural methods. The GLCM (Gray-Level Co-occurrence matrix) method is included in the statistical method, where in statistical calculations, it uses the distribution of gray degrees (histograms) by measuring the level of contrast, granularity, and roughness of an area from the neighboring relations between pixels in the image. This statistical paradigm has unlimited uses, so it is suitable for unstructured natural textures of sub-patterns and rule sets (microstructures).

Artificial Intelligence (Artificial Intelligence or AI) is intelligence that resembles the human brain's decision-making performance. Such systems are generally considered to be computers. Intelligence is created and put into a machine (computer) to do work as humans can. Several fields that use artificial intelligence include expert systems, computer games (games), fuzzy logic, artificial neural networks, and robotics (Ole, 2008).

LVQ is an ANN method for conducting learning in a supervised competitive layer. A competitive layer will automatically learn to classify input vectors. The classes obtained from this competitive layer depend only on the distance between the input vectors. If the two input vectors are close to the same, then the competitive layer will put the two input vectors into the same class. (Kusumadewi, 2004).

# MATERIALS AND METHODS

The data used in this study are MRI images of standard conditions (15 images) and stroke-diagnosed MRI images (57 images). The data was collected from Haji Hospital, Sukolilo, Surabaya, Indonesia.

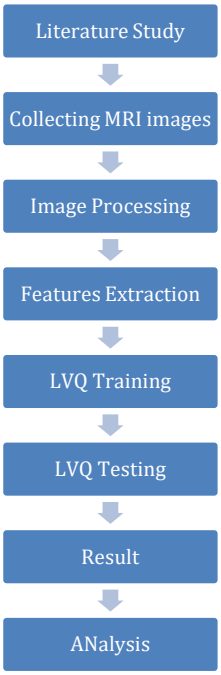



**FIGURE 1.** Flowchart of The proposed Study

**1. Literature Study**

The literature study was carried out by analyzing both internet journals and journals in the reading room of the Faculty of Science and Technology, Airlangga University. Literature studies were also carried out by reading books related to the brain, brain pathology, and MRI at the Medical Library Campus A,

Airlangga University, as well as books and journals related to image processing, ANN LVQ method, and discussing with MRI practitioners at Haji Sukolilo Hospital, Surabaya.

**2. Image Processing**

Image processing will be carried out first by grouping the images according to the category in each folder.

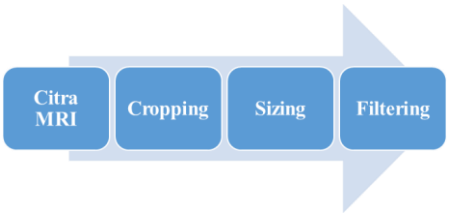



**FIGURE 2.** Image Processing Flowchart

**3. Features Extraction**

Extra image in research using GLCM (Grey Level Co-occurrence Matric) with the features used are: homogeneity, diversity, contrast, variance, and correlation with spatial relationship distance (d) is 1 pixel with 0 direction (Homogeneity, Diversity, Contrast, Correlation)

**4. Design View**

The test window is a valuable program for evaluating previous training results. Sample testing still uses the LVQ method shown in the Sample Testing GUI Design

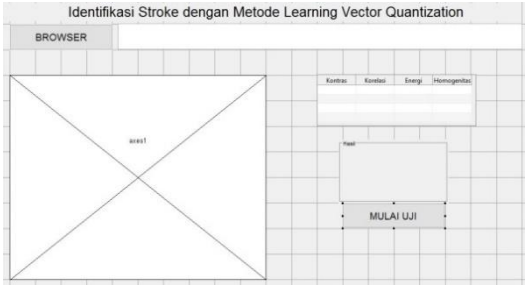



**FIGURE 3.** Sample Testing GUI Design

**5. Training with the LVQ Method**

After determining the matrix value of the GLCM extraction, the next step is to test the input using the LVQ method.

**6. Testing with the LVQ Method**

After training the sample with LVQ, the sample was tested using the same method, namely LVQ.

**7. Training Analysis**

The training data results are in the form of alpha variations, and alpha decreases to the accuracy percentage of each alpha with a decreased alpha.

$$accuracy\ (i) = \frac{\sum j(i) blockage}{n}\ x\ 100\ \%$$

**8. Testing Analysis**

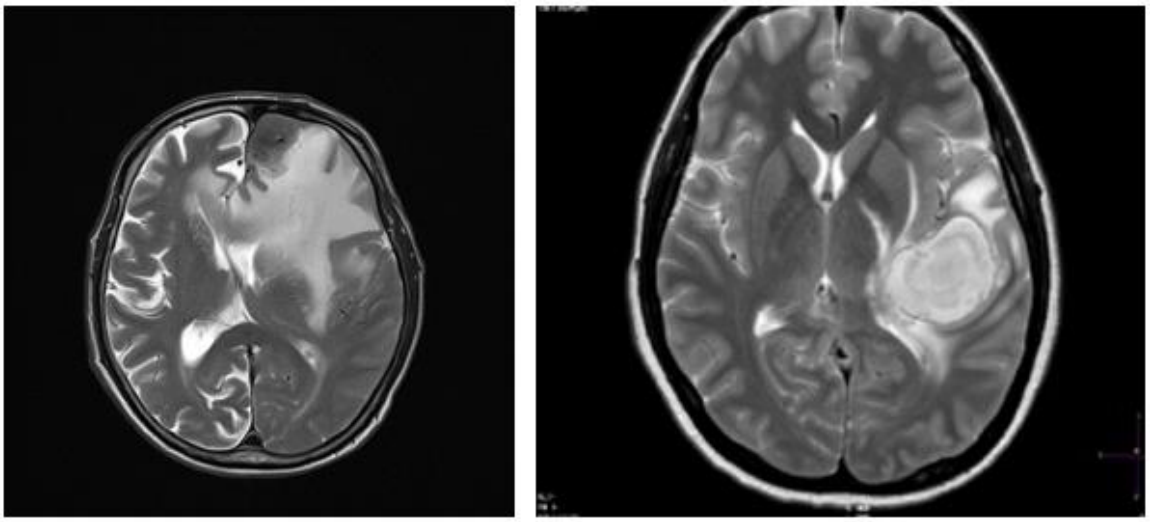
Completed the sample training process to produce the first final weight and the second final weight, which then, in the testing process, becomes the first initial weight and the second initial weight, which then calculates the closest distance to the weight using the formula below (Kusumadewi, 2014)

$$t_1 = \sqrt{\sum (x_i - {w_1}^2)}$$

$$t_2 = \sqrt{\sum (x_i - {w_2}^2)}$$

## RESULTS AND DISCUSSIONS

The application of brain hemorrhage analysis on MRI resembles a brain tumor in terms of shape; it's just a different color, so the image analysis is carried out using a parameter approach by image analysis in brain tumor cases. Shows the difference between the images of a bleeding stroke and a brain tumor, both images obtained using MRI in different patients.



(a) (b)

**FIGURE 4.** (a) brain hemorrhage MRI image (b) brain tumor image

### Learning Vector Quantization Training Process

In the training process, 60 patient images were used, consisting of 50 stroke brain images and 10 standard brain images. The entire image has a BMP file format with image details.

**Table 1**. Image Details

| Attribute | Value |
|---|---|
| Column | 384 |
| Row | 384 |
| Class | Unit8 |
| Image Type | Truecolor |

### Digital Image Processing

BMP files are a type of color format with a proper color type with a bit depth of 24 bits. The median filter cannot process in 24-bit conditions, so it must be changed to 8-bit. One image with 8 bits is gray-scale. Where the image has the intensity of the gray degree (Romli, 2015), the difference in the image is shown in Figure 5.

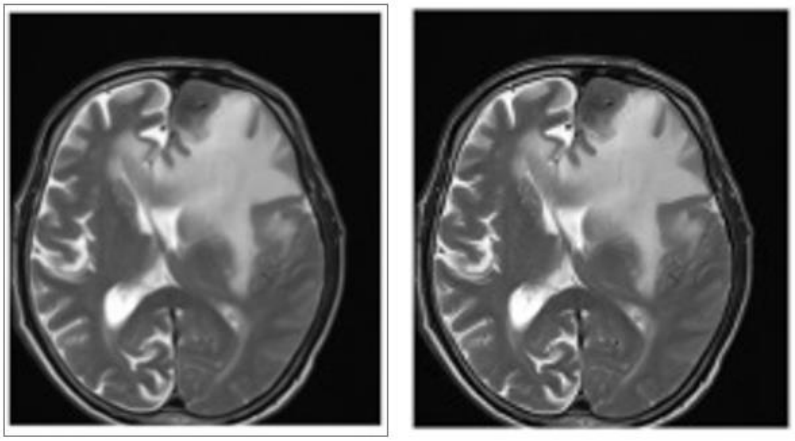



**FIGURE 5.** (a) file in bit depth 24 bit and (b) file in bit depth 8 bit

## Feature extraction with Gray-Level Co-Occurrence Matrix (GLCM)

This study used four parameters out of seven GLCM parameters, namely contrast, correlation, homogeneity, and diversity (Zahoor, 2015). Apart from these four parameters, this study used the same pixel spacing as the same study, namely 0°.

**Table 2.** Four parameters of GLCM

| File Name | Contrast | Correlation | Diversity | Homogeneity |
|---|---|---|---|---|
| Sampel1.BMP | 0,192137 | 0,96624 | 0,379646 | 0,948378 |
| Sampel2.BMP | 0,18168 | 0,965888 | 0,415135 | 0,954494 |

## Results of the Training Data Program

The LVQ data training program that was run resulted in various variations, including sample variations that occupied the initial weight of one (w1) and weight of two (w2). Because the flow of the algorithm from LVQ states that in the data set, the first and second orders act as the initial weights in each existing category, in the study of identifying strokes on MRI images, there are two categories, so the initial weights are the initial weights (w1) representing standard images and the initial weights (w2) representing stroke images.

The initial weight of the training data is the initial determinant of the test data's accuracy level. Therefore, the selection of initial weights must represent the uniformity of all samples below. The selection of the best initial weight is the weight that provides the best accuracy. The best accuracy and value of the two initial and final weights can be seen in the workspace's final table values.

When the program is run, four inputs are used, namely x1 (contrast), x2 (correlation), x3 (diversity), and x4 (homogeneity), with several alpha variations and alpha decreases. The following variations were obtained from previous studies with alpha variations between 0.100 and 0.001 for variations in alpha reduction of 0.010 to 0.500 (Metta, 2009). Whereas for this study, an alpha variation was added between 1 to 1.1, and there was no additional alpha decrease variation. The results of these variations are shown in Table 3.

**Table 3.** Variation of Alpha reduction

| Alpha | Alpha Reduction | Accuracy | Epoch |
|---|---|---|---|
| 0,01 | 0,01 | 74,97% | 230 |
| | 0,1 | 65,50% | 22 |
| | 0,25 | 64% | 9 |
| | 0,5 | 63,80% | 4 |
| 0.001 | 0,01 | 46,50% | 78 |
| | 0,1 | 46,50% | 49 |
| | 0,25 | 46,50% | 21 |
| | 0,5 | 46,50% | 2 |

From the alpha variations and alpha decreases table, the best accuracy is 74.97% with an error of 25.03% at an alpha of 0.01 with an alpha decrease of 0.010. This result is still less than the results of previous research by Metta (2009) which has the best accuracy of 80%. From the Matlab Document reference, the default value for alpha is 0.1, while the results in the table above for an alpha value of 0.1 produce an accuracy of 72.50% for each given alpha decrease. So that in the training process, the data is not suitable for using the default alpha value from the Matlab Document Table 4 shows the first weight (W1), the initial second weight (W2) with the first weight (W2), and the second weight (W2) the final result obtained from the iteration (epoch) reaches a maximum of 230.

**Table 4.** Parameters with different weight

| W | Contrast | Correlation | Diversity | Homogeneity |
|---|---|---|---|---|
| **Initial W1 and W2** | | | | |
| $W_1$ | 0,181680 | 0,965888 | 0,415135 | 0,954494 |
| $W_2$ | 0,189431 | 0,971257 | 0,427128 | 0,949997 |
| **Final W1 and W2** | | | | |
| $W_1$ | 0,181542 | 0,934099 | 0,404993 | 0,952417 |
| $W_2$ | 0,182578 | 0,972686 | 0,424034 | 0,951037 |

After getting the final W1 and W2 weights, these weights become the initial W1 and W2 for the sample testing process and the best parameters in the LVQ test because the final W1 and W2 values are the final values of the highest accuracy in the data training process.

## Learning Vector Quantization (LVQ) Testing Process

In the sample testing process, there are 14 test data consisting of images that have not been categorized as normal or stroke. The existence of a random sample test is intended to determine the performance of the program in stroke identification. The test results for the 14 test data are attached on the attachment page.

The sample testing process is different from the training process. In the testing process, there is no decrease in alpha, so the alpha is constant, and there is no renewal of the latest weights.

## Graphical User Interface (GUI) Testing

Figure 6 shows the graphical user interface for LVQ testing.

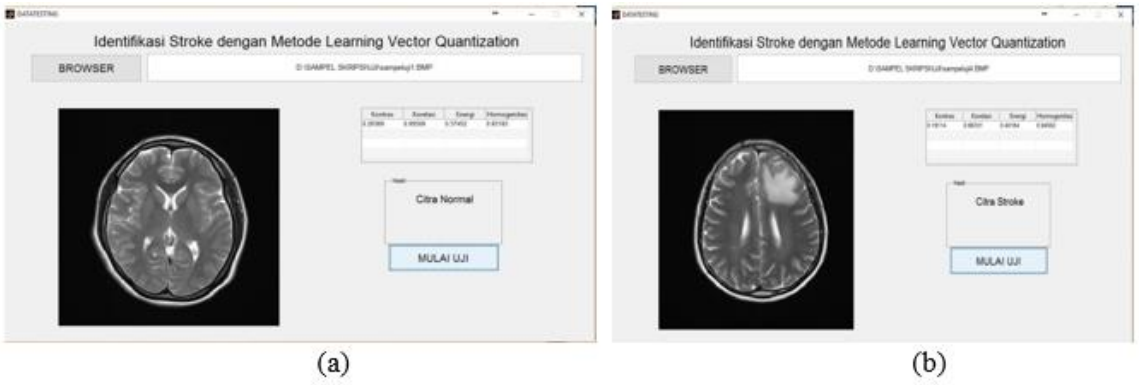



**FIGURE 6.** (a) file in bit depth 24 bit and (b) file in bit depth 8 bit

## Test Program Results

The test samples given to the testing program were 14 normal and stroke images, with the results of the program testing attached in the Appendix of Program Testing Results. From the results of the program testing, there are several variables, namely x1, which is contrast; x2, which is a correlation; x3, which is diversity; and x4, which is homogeneity. An arithmetic operation from the four x values is given using equation 3.2 to produce the t1 value and equation 3.3 to produce the t2 value. The values of t1 and t2 are the weight distances, with t1 being the normal image category weight distances and t2 being the stroke image category weight distances. Identification decisions are at their closest. If it is closest to t1, then the sample will be classified as a normal image and vice versa.

## CONCLUSION

From the research that has been done, it can be concluded that the magnitude of the error value in the program is due to the lack of data used in training so that the training data is less diverse, causing when the program is tested, the program built is less training and provides an accuracy of 78.60%. However, referring to the original purpose of making the program for secondary opinion, practitioners can use the program..

## REFERENCES

1. Fadli.SM..2013. *Pengetahuan dan ekspresi emosi keluarga serta frekuensi kekambuhan penderita skizofrenia.*program studi megister ilmu kesehatan masyarakat sekolah tinggi ilmu kesehatan hang tuah. pekanbaru
2. Juniarti.S. 2008 *Buletin Jendela Data Informasi Dan Kesehatan.* Kementrian Kesehatan RI. Jakarta
3. Kusumadewi.S.2004. *Artificial Intelligence (Teknik Dan Aplikasinya)*.Graha Ilmu. Jakarta
4. Matlab Documentation. The Mathworks, Inc
5. Metta.S.2009.*Identifikasi citra CT Scan kanker paru – paru dengan Metode Learning Vector Quantization*. Fakultas Sains dan Teknologi Universitas Airlangga Surabaya. Surabaya
6. Notosiswoyo.M.2004.*Pemanfaatan Magnetic Resonance Imagin (MRI) Sebagai Sarana Diagnosis Pasien.*Media Litbang Kesehatan Volume XIV Nomor 3
7. Ole.Y..2008.*Artificial Intelegence*.jurusan informatika fakultas teknik informatika universitas Sumatra utara. Sumatra utara
8. Prionika.D.2012. *Kecerdasan Buatan Untuk Pemula.*Jogjakarta
9. Pupitaningrum.D..2004 *pengantar jaringan syaraf tiruan.*fakultas teknik. jurusan teknologi informasi. universitas semarang
10. Riyadina.W.2013. *determinan penyakit stroke.*badan penelitian dan pengembangan kesehatan kementrian kesehatan republik Indonesia
11. Samsudin.Dr.H.2009 *Buletin Jendela Data Informasi Dan Kesehatan.* Kementrian Kesehatan RI Jakarta
12. Tatang.2015.*Penyebab Terjadinya Kerusakan Otak Manusia.* Tatangsma.Com/2015/03/Penyebab-Terjadinya-Kerusakan-Otak-Manusia.Html

13. Tjahjadi.Hendrana.2014 *Magnetic Resonance Imagin*.Jurusan Teknobiomedik.Program Pasca Sarjana.Universitas Indonesia
14. Valadka.Ab..Gonipath Sp.. *Et Al*.1996.Relationship Of Brain Tissue Po2 To Outcome After Severe Head Injury. Crift Care Med
15. Wismanto..2015 *Penunjang Diagnosis Fisioterapi*.Fakultas Kedokteran.Universitas Airlangga
16. Zahoor.Arbaaz.2015 *Computer Aided Diagnosis of Stroke from Brain CT Images*.M. Tech Student.SJCE Mysore