DOKUZ EYLÜL UNIVERSITY
GRADUATE SCHOOL OF NATURAL
AND APPLIED SCIENCES

MS SEGMENTATION BY USING MRI IMAGES

by
Şule ÜŞÜMEZOĞLU

June, 2007
İZMİR

MS SEGMENTATION
BY USING MRI IMAGES

A Thesis Submitted to the
Graduate School of Natural and Applied Sciences of Dokuz Eylül University
In Partial Fulfillment of the Requirements for the Degree of Master of Science
in
Electrical and Electronics Engineering

by
Şule ÜŞÜMEZOĞLU

June, 2007
İZMİR

M.Sc THESIS EXAMINATION RESULT FORM

We have read the thesis entitled **"MS SEGMENTATION BY USING MRI IMAGES"** completed by **ŞULE ÜŞÜMEZOĞLU** under supervision of **ASST. PROF. DR. AHMET ÖZKURT** and we certify that in our opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

Ass. Prof. Dr. Ahmet ÖZKURT

Supervisor

Asst. Prof. Dr. Radosveta Ivanova Sokullu
Jury Member

Asst. Prof. Dr. Mehmet Kuntalp
Jury Member

Prof.Dr. Cahit HELVACI
Director
Graduate School of Natural and Applied Sciences

ACKNOWLEDGEMENTS

I would like to thank to my supervisor Asst. Prof. Dr. Ahmet Özkurt because of his helps, suggestions and support for the researching and writing of this thesis. I also want to thank to Asst. Prof. Dr. Nalan Özkurt because of her kindly support. I would like to thank to my dear friend Taner Akkan, because of his encouragement who is always belived me. I want to present my thanks to the staff of the Graduate School of Natural and Applied Science, especially to Mrs. Filiz Gürsan.

Finally, thanks for their valuable helps and patience, my mother Şahamet İsmet Altuğ, my husband İsmail Üşümezoğlu and my unique son Cemre Burak. I am always indebted to their unconditional love.

Şule ÜŞÜMEZOĞLU

MS SEGMENTATION BY USING MRI IMAGES

ABSTRACT

Multiple Sclerosis (MS) is a disease of the central nervous system, brain and spinal cord. The inherent heterogeneity of MS lesions, which is reflected on MR images, causes some of the difficulties encountered while attempting to separate lesions from healthy brain tissue and cerebro-spinal fluid. Automated segmentation of the lesions is useful for trace and diagnosis of MS. In this study, different tissue classification techniques are compared for segmentation of the MS lesions. Artificial neural networks are considered as supervised segmentation techniques and clustering algorithms are considered as unsupervised techniques. Fuzzy C-Means, K-Means and K-medoid clustering techniques with Principle Component Analysis, Independent Component Analysis and Multi Layer Back Propagation Neural Network algorithms are compared and results are evaluated.

Keywords: Multiple Sclerosis, Magnetic Resonance Imaging, Segmentation, Artificial Neural Network, Fuzzy C-Means, K-Means, Clustering, Principle Component Analysis (PCA), Independent Component Analysis (ICA).

MR GÖRÜNTÜLERİ KULLANARAK MS BÖLÜTLEMESİ

ÖZ

Multiple Sclerosis (MS), merkezi sinir sistemi, beyin ve omuriliği etkileyen bir hastalıktır. MS lezyonlarının içsel heterojen yapısı MR görüntülerine yansır ve lezyonların sağlıklı beyin dokusundan ayrılması sırasında zorluklarla karşılaşılır. Lezyonların otomatik bölütlenmesi MS'in tanımlanması ve izlenmesi açısından gereklidir. Bu çalışmada, MS lezyonlarının bölütlenmesi için farklı doku sınıflandırma teknikleri karşılaştırıldı. Yapay sinir ağları, güdümlü yöntem olarak dikkate alındı. Bağımsız bileşen analizi (ICA) ve temel bileşen analizi (PCA) ile birleştirilmiş fuzzy c-means, k-means ve k-medoid kümeleme yöntemleri ile çok seviyeli geri yayılım sinir ağları algoritmaları karşılaştırıldı ve sonuçlar değerlendirildi.

Anahtar Sözcükler: Multiple Sclerosis, Manyetik Rezonans Görüntüleme, Bölütleme, Yapay Sinir Ağları, Fuzzy C-Means, K-Means, Kümeleme, Temel Bileşen Analizi (PCA), Bağımsız Bileşen Analizi (ICA).

CONTENTS

Page

THESIS EXAMINATION RESULT FORM…… ii
ACKNOWLEDGEMENTS…….. iii
ABSTRACT…….. iv
ÖZ…….. v

CHAPTER ONE-INTRODUCTION……..…… 1

1.1 Magnetic Resonance Imaging (MRI)…….. 1

1.2 Multiple Sclerorsis (MS)..……... 1
1.2.1 Signs and Symptomps of MS…….. 2
1.2.2 Diagnosis of MS……. 2
1.2.3 Disease Course and Clinical Subtypes of MS…….. 4
1.2.3.1 Relapsing-remitting…….. 4
1.2.3.2 Secondary progressive…….. 5
1.2.3.3 Primary progressive…….. 5
1.2.3.4 Progressive relapsing…… 5
1.2.4 Factors triggering a relapse of MS…… 6
1.2.5 Pathophysiology of MS…… 7
1.2.6 Causes of MS……... 8
1.2.6.1 Environmental…….. 8
1.2.6.2 Genetic……. 10
1.2.7 Epidemiology…….. 12

1.3 Detection of MS with MRI…….…… 13

1.4 Automated Detection of Multiple Sclerosis (MS) Brain Lesions in Magnetic Resonance Imaging (MRI)…….. 14

CHAPTER TWO- MAGNETIC RESONANCE IMAGING……16

2.1 Magnetic Resonance Imaging……..16

CHAPTER THREE-IMAGE CLUSTERING METHODS……... 24

3.1 Unsupervised Learning And Clustering…….. 24

3.2 Clustering Algorithms…… 25
3.2.1 K-Means And K-Medoid Algorithms……..25
3.2.2 Fuzzy C-Means Algorithm…….. 26

3.3 Component Analysis……28
3.3.1 Principle Component Analysis (PCA)…….28
3.3.2 Non-Linear Component Analysis (NLCA)…… 31
3.3.3 Independent Component Analysis (ICA)…… 31

3.4 Supervised Learning…….……... 32
3.4.1 Multilayer Neural Networks……32
3.4.1.1 A Neuron Model…….. 33
3.4.1.2 Backpropagation Learning Algorithm (MLP NN)……..... 34

CHAPTER FOUR-MS SEGMENTATION……...36

4.1 The Data…….. ……36

4.2 Segmentation Methods…… 37

4.3 Input Data…… 38

4.4 Segmentation With Clustering Techniques……. 38

4.4.1 Fuzzy C-Means Clustering…….. 38
4.4.2 K-Means Clustering……42
4.4.3 K-Medoid Clustering……... 46

4.5 Principle Component Analysis (PCA)…… 50
4.5.1 FCM With PCA…….. 53
4.5.2 K-Means With PCA…… 58
4.5.3 K-Medoid With PCA…….. 63

4.6 Independent Component Analysis (ICA)……...…. 68

4.7 Segmentation With Neural Networks……. 73
4.7.1 Segmentation With The Feedforward Multilayer Perceptron Algorithm……...……. 73
4.7.2 Application of Feedforward Multilayer Perceptron Neural Network……..…… 74

4.8 Comparison of Techniques……..….. 79
4.8.1 Comparison of Elapsed Times…… 81
4.8.2 Comparison of Percentage of the True Pixels…….. 81
4.8.3 Comparison of the False Positive Pixel Numbers…….. 82
4.8.4 Comparison of the False Negative Pixel Numbers……. 82
4.8.5 Sensitivity…….82
4.8.5.1 Sensitivity of clustering methods…….83
4.8.5.2 Sensitivity of independent component analysis method…….83
4.8.5.3 Sensitivity of multilayer backpropagation neural network method…….……..……..83
4.8.6 Specificity……84
4.8.6.1 Specificity of clustering methods……...84
4.8.6.2 Specificity of independent component analysis method……. 85
4.8.6.3 Specificity of multilayer backpropagation neural network method……. 85

4.9 Result……..…… 86

4.10 GUI…….86

CHAPTER FIVE-CONCLUSION…… 89

5.1 Future Work…… ……90

REFERENCES……92

APPENDIX…..…… 97

CHAPTER ONE
INTRODUCTION

Multiple Sclerosis (MS) is a disease of the central nervous system, the brain and spinal cord. Magnetic resonance imaging is very important for radiologists for detecting and seperating lesion areas. The inherent heterogeneity of MS lesions, which is reflected on MR images, causes some of the difficulties encountered while attempting to separate lesions from healthy brain tissue and cerebro-spinal fluid. Automated segmentation of the lesions is useful for trace and diagnosis of MS.

1.1 Magnetic Resonance Imaging (MRI)

Magnetic Resonance Imaging is important technique for diagnosis diseases and soft tissue contrast is ideal modality for investigating central nervous system disorders. Magnetic Resonance Imaging is a non-invasive medical technique which through the production of series of cross-sectional images, allows the visualization of internal anatomical structures of the body. MRI is typically used qualitatively, with radiologists inspecting the acquired images for tumors, lesions, and other abnormalities. Quantitative estimation of tissue volumes provides important information about the natural progression of disease and helps to evaluate efficiency of threapeutic intervension (Hornac, J.P.,2007).

1.2 Multiple Sclerosis (MS)

Multiple Sclerosis (MS) is a disease of the central nervous system, the brain and spinal cord. It is a debilitating and progressive disease which may result in a variety of symptoms from blurred vision to severe muscle weakness and degradation, depending on the area of the central nervous system which is affected. Multiple Sclerosis is caused by a breakdown in the myelin sheath, a soft, white, fatty material which insulates the neurons of the central nervous system and provides for the rapid transmission of nerve impulses along the nerve fibers of the brain and spinal cord. MS lesions go through various hist-pathological stages and are heterogeneous in their

presentation. Some chronic lesions may contain “scar” tissue composed of astrocytic fibers (Pu Ma, 2001).

1.2.1 Signs and Symptomps of MS

MS can cause a variety of symptoms, including changes in sensation, muscle weakness, abnormal muscle spasms, or difficulty in moving, difficulties with coordination and balance, problems in speech or swallowing, visual problems, fatigue and acute or chronic pain syndromes, bladder and bowel difficulties, cognitive impairment, or emotional symptomatology (http://en.wikipedia.org/ wiki/Multiple_sclerosis#_note-2, June 2007) .

The main clinical measure of disability progression and severity of the symptoms is the Expanded Disability Status Scale or EDSS (Kurtzke, J.F., 1983).

The initial attacks are often transient, mild, and self-limited. They often do not prompt a health care visit and sometimes are only identified in retrospect once the diagnosis has been made based on further attacks. The most common initial symptoms reported are: Changes in sensation in the arms, legs or face (33%), complete or partial vision loss (16%), weakness (13%), double vision (7%), unsteadiness when walking (5%), and balance problems (3%) but many rare initial symptoms have been reported such as aphasia or psychosis (Navarro, S., Mondéjar-Marín, B., 2005), (Jongen, P., 2006). Fifteen percent of individuals have multiple symptoms when they first seek medical attention (Paty, D., Studney D.,1994). For some people the initial MS attack is preceded by infection, trauma, or strenuous physical effort.

1.2.2 Diagnosis of MS

Multiple sclerosis is difficult to diagnose in its early stages. In fact, definite diagnosis of MS cannot be made until there is evidence of at least two anatomically separate demyelinating events occurring at least thirty days apart (http:// en.

wikipedia.org/ wiki/Multiple_sclerosis#_note-2, June 2007) .

Historically different criteria were used. The Schumacher criteria and Poser criteria were both popular. Currently, McDonald criteria represents international efforts to standardize the diagnosis of MS using clinical data, laboratory data, and radiologic data (McDonald, W., Compston A., 2001).

Clinical data alone may be sufficient for a diagnosis of MS. If an individual has suffered two separate episodes of neurologic symptoms characteristic of MS, and the individual also has consistent abnormalities on physical examination, a diagnosis of MS can be made with no further testing. Since some people with MS seek medical attention after only one attack, other testing may hasten the diagnosis and allow earlier initiation of therapy.

Magnetic resonance imaging (MRI) of the brain and spine is often used to evaluate individuals with suspected MS. MRI shows areas of demyelination as bright lesions on T2-weighted images or FLAIR (fluid attenuated inversion recovery) sequences. Gadolinium contrast is used to demonstrate active plaques on T1-weighted images. Because MRI can reveal lesions which occurred previously but produced no clinical symptoms, it can provide the evidence of chronicity needed for a definite diagnosis of MS.

Testing of cerebrospinal fluid (CSF) can provide evidence of chronic inflammation of the central nervous system. The CSF is tested for oligoclonal bands, which are immunoglobulins found in 85% to 95% of people with definite MS (Rudick, R., Whitaker, J., 1987). Combined with MRI and clinical data, the presence of oligoclonal bands can help make a definite diagnosis of MS. Lumbar puncture is the procedure used to collect a sample of CSF.

The brain of a person with MS often responds less actively to stimulation of the optic nerve and sensory nerves. These brain responses can be examined using visual

evoked potentials (VEPs) and somatosensory evoked potentials (SEPs). Decreased activity on either test can reveal demyelination which may be otherwise asymptomatic. Along with other data, these exams can help find the widespread nerve involvement required for a definite diagnosis of MS (Gronseth, G., Ashman, E., 2000).

Another test which may become important in the future is measurement of antibodies against myelin proteins such as myelin oligodendrocyte glycoprotein (MOG) and myelin basic protein (MBP). As of 2007, however, there is no established role for these tests in diagnosing MS (http:// en.wikipedia.org/ wiki/Multiple_sclerosis#_note-2, June 2007).

The signs and symptoms of MS can be similar to other medical problems, such as neuromyelitis optica, stroke, brain inflammation, infections such as Lyme disease (which can produce identical MRI lesions and CSF abnormalities), tumors, and other autoimmune problems, such as lupus. Additional testing may be needed to help distinguish MS from these other problems (Garcia-Monco, J., Miro, J., 1990), (Hansen, K., Cruz, M., 1990), (Schluesener, H., Martin, R., 1989), (Kohler, J., Kern, U., 1988).

1.2.3 Disease Course and Clinical Subtypes of MS

The course of MS is difficult to predict, and the disease may at times either lie dormant or progress steadily. Several subtypes, or patterns of progression, have been described. Subtypes use the past course of the disease in an attempt to predict the future course. Subtypes are important not only for prognosis but also for therapeutic decisions. In 1996 the United States National Multiple Sclerosis Society standardized the following four subtype definitions (Lublin, F., Reingold, S., 1996).

1.2.3.1 Relapsing-remitting

Relapsing-remitting describes the initial course of 85% to 90% of individuals with MS. This subtype is characterized by unpredictable attacks (relapses) followed by

periods of months to years of relative quiet (remission) with no new signs of disease activity. Deficits suffered during the attacks may either resolve or may be permanent. When deficits always resolve between attacks, this is referred to as "benign" MS.

1.2.3.2 Secondary progressive

Secondary progressive describes around 80% of those with initial relapsing-remitting MS, who then begin to have neurologic decline between their acute attacks without any definite periods of remission. This decline may include new neurologic symptoms, worsening cognitive function, or other deficits. Secondary progressive is the most common type of MS and causes the greatest amount of disability.

1.2.3.3 Primary progressive

Primary progressive describes the approximately 10% of individuals who never have remission after their initial MS symptoms. Decline occurs continuously without clear attacks. The primary progressive subtype tends to affect people who are older at disease onset.

1.2.3.4 Progressive relapsing

Progressive relapsing describes those individuals who, from the onset of their MS, have a steady neurologic decline but also suffer superimposed attacks; and is the least common of all subtypes.

Special cases of the disease with non-standard behavior have also been described although many researchers believe they are different diseases. These cases are sometimes referred to as borderline forms of multiple sclerosis and are Neuromyelitis optica (NMO), Balo concentric sclerosis, Schilder's diffuse sclerosis and Marburg multiple sclerosis (Fontaine, B., 2001).

1.2.4 Factors triggering a relapse of MS

Multiple sclerosis relapses are often unpredictable and can occur without warning with no obvious inciting factors. Some attacks, however, are preceded by common triggers. In general, relapses occur more frequently during spring and summer than during autumn and winter. Infections, such as the common cold, influenza, and gastroenteritis, increase the risk for a relapse (Confavreux, C., 2002). Emotional and physical stress may also trigger an attack, as can severe illness of any kind.

Statistically, there is no good evidence that either trauma or surgery trigger relapses (Buljevac D., Hop W., 2003), (Brown R., Tennant C., 2006). People with MS can participate in sports, but they should probably avoid extremely strenuous exertion, such as marathon running. Heat can transiently increase symptoms, which is known as Uhthoff's phenomenon. This is why some people with MS avoid saunas or even hot showers. However, heat is not an established trigger of relapses (Tataru N., Vidal C., 2003)

Pregnancy can directly affect the susceptibility for relapse. The last three months of pregnancy offer a natural protection against relapses. However, during the first few months after delivery, the risk for a relapse is increased 20%–40%. Pregnancy does not seem to influence long-term disability. Children born to mothers with MS are not at increased risk for birth defects or other problems (Worthington J., Jones R., 1994).

Many potential triggers have been examined and found not to influence relapse rates in MS. Influenza vaccination is safe, does not trigger relapses, and can therefore be recommended for people with MS. There is also no evidence that hepatitis B, varicella, tetanus, or Bacille Calmette-Guerin (BCG immunization for tuberculosis) increases the risk for relapse (Confavreux C., Suissa S., 2001).

1.2.5 Pathophysiology of MS

Although much is known about how multiple sclerosis causes damage, the reasons why multiple sclerosis occurs are not known (http://en.wikipedia.org/ wiki/ Multiple_sclerosis#_note-2, June 2007).

Multiple sclerosis is a disease in which the myelin (a fatty substance which covers the axons of nerve cells) degenerates. According to the view of most researchers, a special subset of lymphocytes, called T cells, plays a key role in the development of MS.

According to a strictly immunological explanation of MS, the inflammatory processes triggered by the T cells create leaks in the blood-brain barrier (a capillary system that should prevent entrance of T-cells into the nervous system). These leaks, in turn, cause a number of other damaging effects such as swelling, activation of macrophages, and more activation of cytokines and other destructive proteins such as matrix metalloproteinases. A deficiency of uric acid has been implicated in this process (Rentzos, M., Nikolaou, C., 2006).

In a person with MS, these lymphocytes recognize myelin as foreign and attack it as if it were an invading virus. That triggers inflammatory processes, stimulating other immune cells and soluble factors like cytokines and antibodies.

It is known that a repair process, called remyelination, takes place in early phases of the disease, but the oligodendrocytes that originally formed a myelin sheath cannot completely rebuild a destroyed myelin sheath. The newly-formed myelin sheaths are thinner and often not as effective as the original ones. Repeated attacks lead to successively fewer effective remyelinations, until a scar-like plaque is built up around the damaged axons, according to four different damage patterns (Lucchinetti, C., Bruck, W., 2000). The central nervous system should be able to recruit oligodendrocyte stem cells capable of turning into mature myelinating

oligodendrocytes, but it is suspected that something inhibits stem cells in affected areas.

Also the axons are damaged by the attacks (Pascual A., Martínez-Bisbal M., 2007). Often, the brain is able to compensate for some of this damage, due to an ability called neuroplasticity. MS symptoms develop as the cumulative result of multiple lesions in the brain and spinal cord. This is why symptoms can vary greatly between different individuals, depending on where their lesions occur.

1.2.6 Causes of MS

Although many risk factors for multiple sclerosis have been identified, no definitive cause has been found. MS likely occurs as a result of some combination of both environmental and genetic factors. Various theories try to combine the known data into plausible explanations. Although most accept an autoimmune explanation, several theories suggest that MS is an appropriate immune response to an underlying condition. The need for alternative theories is supported by the poor results of present therapies, since autoimmune theory predicted greater success (Behan, P., Chaudhuri, A., 2002), (Chaudhuri A, Behan P., 2004), (Altmann, D., 2005).

1.2.6.1 Environmental

The most popular hypothesis is that a viral infection or retroviral reactivation primes a susceptible immune system for an abnormal reaction later in life. On a molecular level, this might occur if there is a structural similarity between the infectious virus and some component of the central nervous system, leading to eventual confusion in the immune system.

Since MS seems to be more common in people who live farther from the equator, another theory proposes that decreased sunlight exposure (Ponsonby, A., Dwyer, T., 2003) and possibly decreased vitamin D production may help cause MS. This theory is bolstered by recent research into the biochemistry of vitamin D, which has shown that it is an important immune system regulator. A large, 2006 study by the Harvard

School of Public Health, reported evidence of a link between Vitamin D deficiency and the onset of multiple sclerosis (Munger, K., Levin, L., 2006). Other data comes from an 2007 study which concluded that sun exposure during childhood reduces the risk of suffering MS, while controlling for genetic factors (Talat, I., 2007).

Other theories, noting that MS is less common in children with siblings, suggest that less exposure to illness in childhood leads to an immune system which is not primed to fight infection and is thus more likely to attack the body. One explanation for this would be an imbalance between the Th1 type of helper T-cells, which fight infection, and the Th2 type, which are more active in allergy and more likely to attack the body.

Other theories describe MS as an immune response to a chronic infection. The association of MS with the Epstein-Barr virus suggests a potential viral contribution in at least some individuals (Levin, L., Munger K., 2005). Still others believe that MS may sometimes result from a chronic infection with spirochetal bacteria, a hypothesis supported by research in which cystic forms were isolated from the cerebrospinal fluid of all MS patients in a small study (Brorson, O., Brorson S., 2001). When the cysts were cultured, propagating spirochetes emerged. Another bacterium that has been implicated in MS is Chlamydophila pneumoniae, it or its DNA has been found in the cerebrospinal fluid of MS patients by several research laboratories, with one study finding that the oligoclonal bands of 14 of the 17 MS patients studied consisted largely of antibodies to Chlamydophila antigens (Yao, S., Stratton, C., 2001).

Severe stress may also be a factor a large study in Denmark found that parents who had lost a child unexpectedly were 50% more likely to develop MS than parents who had not (Li, J., Johansen C., 2004). Smoking has also been shown to be an independent risk factor for developing MS (Franklin, G., Nelson, L., 2003).

1.2.6.2 Genetic

MS is not considered a hereditary disease. However, increasing scientific evidence suggests that genetics may play a role in determining a person's susceptibility to MS.

Some populations, such as the Roma, Inuit, and Bantus, rarely if ever get MS. The indigenous peoples of the Americas and Asians have very low incidence rates.

In the population at large, the chance of developing MS is less than a tenth of one percent. However, if one person in a family has MS, that person's first-degree relatives (parents, children, and siblings) have a one to three percent chance of getting the disease.

For identical twins, the likelihood that the second twin may develop MS if the first twin does is about 30%, for fraternal twins (who do not inherit identical gene pools), the likelihood is closer to that for non-twin siblings, or about 4%. The fact that the rate for identical twins both developing MS is significantly less than 100% suggests that the disease is not entirely genetically controlled. Some (but definitely not all) of this effect may be due to shared exposure to something in the environment, or to the fact that some people with MS lesions remain essentially asymptomatic throughout their lives.

Further indications that more than one gene is involved in MS susceptibility comes from studies of families in which more than one member has MS. Several research teams found that people with MS inherit certain regions on individual genes more frequently than people without MS. Of particular interest is the human leukocyte antigen (HLA) or major histocompatibility complex region on chromosome 6. HLAs are genetically determined proteins that influence the immune system. However, there are other genes in this region which are not related to the immune system.

The HLA patterns of MS patients tend to be different from those of people without the disease. Investigations in northern Europe and America have detected three HLAs that are more prevalent in people with MS than in the general population. Studies of American MS patients have shown that people with MS also tend to exhibit these HLAs in combination that is, they have more than one of the three HLAs more frequently than the rest of the population. Furthermore, there is evidence that different combinations of the HLAs may correspond to variations in disease severity and progression.

A large study examining 334,923 single nucleotide polymorphisms (small variations in genes) in 931 families showed that apart from HLA-DRA there were two genes in which polymorphims strongly predicted MS; these were the IL2RA (a subunit of the receptor for interleukin 2) and the IL7RA (idem for interleukin 7) genes. Mutations in these genes were already known to be associated with diabetes mellitus type 1 and other autoimmune conditions; the findings therefore support the notion that MS is an autoimmune disease (http://content.nejm.org/cgi/ content/ full/NEJMoa073493v1, June 2007).

Studies of families with multiple cases of MS and research comparing proteins expressed in humans with MS to those of mice with EAE suggest that another area related to MS susceptibility may be located on chromosome 5. Other regions on chromosomes 2, 3, 7, 11, 17, 19, and X have also been identified as possibly containing genes involved in the development of MS.

These studies strengthen the theory that MS is the result of a number of factors rather than a single gene or other agent. Development of MS is likely to be influenced by the interactions of a number of genes, each of which (individually) has only a modest effect. Additional studies are needed to specifically pinpoint which genes are involved, determine their function, and learn how each gene's interactions with other genes and with the environment make an individual susceptible to MS.

1.2.7 Epidemiology

In northern Europe, continental North America, and Australasia, about one of every 1000 citizens suffers from multiple sclerosis, whereas in the Arabian peninsula, Asia, and continental South America, the frequency is much lower. In sub-Saharan Africa, MS is extremely rare. With important exceptions, there is a north-to-south gradient in the northern hemisphere and a south-to-north gradient in the southern hemisphere, with MS being much less common in people living near the equator (http://www.fedem.org /revista/ n16/kurtzkeing.htm, June 2007). Climate, diet, geomagnetism, toxins, sunlight exposure, genetic factors, and infectious diseases have all been discussed as possible reasons for these regional differences. Environmental factors during childhood may play an important role in the development of MS later in life. This idea is based on several studies of migrants showing that if migration occurs before the age of fifteen, the migrant acquires the new region's susceptibility to MS. If migration takes place after age fifteen, the migrant keeps the susceptibility of his home country (Marrie, R., 2004).

MS occurs mainly in Caucasians. It is twentyfold lower in the Inuit people of Canada than in other Canadians living in the same region. It is also rare in the Native American tribes of North America, Australian Aborigines and the Maori of New Zealand. Scotland appears to have the highest rate of MS in the world (Rothwell, P., Charlton, D., 1998). The reasons for this are unknown. These few examples point out that either genetic background or lifestyle and cultural factors play an important role in the development of MS.

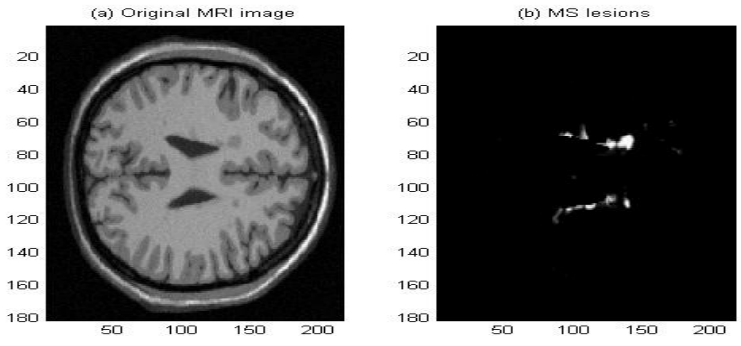
As observed in many autoimmune disorders, MS is more common in females than males; the mean sex ratio is about two females for every male. In children (who rarely develop MS) the sex ratio may reach three females for each male. In people over age fifty, MS affects males and females equally. Onset of symptoms usually occurs between fifteen to forty years of age, rarely before age fifteen or after age sixty.

As previously discussed, there is a genetic component to MS. On average one of every 25 siblings of individuals with MS will also develop MS. Almost half of the identical twins of MS-affected individuals will develop MS, but only one of twenty fraternal twins. If one parent is affected by MS, each child has a risk of only about one in forty of developing MS later in life (Sadovnick, A., Ebers, G., 1996).

Finally, it is important to remark that advances in the study of related diseases have shown that some cases formerly considered MS are not MS at all. In fact, all the studies before 2004 can be affected by the impossibility to distinguish MS and Devic's disease (NMO) reliably before this date. The error can be important in some areas, and is considered to be 30% in Japan (Weinshenker, B., 2005).

1.3 Detection of MS with MRI

MRI detects not only these chronic astrocytic lesions but also the acute edematous patches, with varying amounts of demyelination and inflammatory cell infiltration, typical of newer lesions (Hadjiprocopic, A., 2003). These lesional components (referred to as MS plaques) are also readily observable post-mortem. The inherent heterogeneity of the lesions, which is reflected on MR images causes some of the difficulties encountered while attempting to separate lesions from healthy brain tissue and cerebro-spinal fluid compartments using segmentation approaches based on signal intensity characteristics alone. Fig. 1.1 shows MS plaques in a typical slice from a proton density (T1) weighted MRI image.



(a) (b)

Figure 1.1 (a) MRI image with (b) MS lesions
BrainWeb: Simulated Brain Database, http://www.bic.mni.mcgill.ca/brainweb/

Fig. l.1 (a) shows the image itself while Fig. 1.1 (b) has been overlaid the regions of interest indicating the presence of MS lesions. Note how the lesions are detected almost invariably in the white matter of the brain. Observation of the characteristic plaques over time in MS patients shows how the demyelination of nerves is a reversible and recurrent process in these patients.

In a majority of patients, over time with and without treatment, MS lesions grow and shrink and some lesions may show confluence with others. It is important to be able to quantitatively assess the numbers and sizes of MS lesions in patients undergoing therapy for the disease, and several large scales clinical trials are under way to determine which quantitative measures obtained from the MR images are most useful for assessing the impact of drugs on the disease. Fig. l.1 (b) shows the typical manner in which lesions are quantified in many centers. Much recent research in MS using MRI has been directed toward automating the quantitation of MS lesions in the brain (Narayana P.A.,2004).

1.4 Automated Detection of Multiple Sclerosis (MS) Brain Lesions in Magnetic Resonance Imaging (MRI)

Artificial intelligence techniques of machine learning, pattern recognition and the use of domain knowledge were employed in the segmentation or automated detection of Multiple Sclerosis (MS) lesions in Magnetic Resonance Images of the human brain. Segmentation enables the quantitative measurement of MS lesion volumes, making it important for the study of the disease, the evaluation of drug treatments, and MS patient follow-up. Manual segmentation of MS lesions is tedious, timeconsuming, and difficult. Magnetic Resonance Images contain noise and artifacts which can cause regions of hyperintensity similar to that of MS lesions. Hence, purely data driven techniques of automated segmentation such as intensity thresholding, edge-detection, and multispectral feature-space classification tend to produce a large number of false positive lesions even when images have been preprocessed to reduce impurities.

We can consider the automated segmentation methods as supervised and unsupervised techniques. Supervised techniques are needed to training data to classify the clusters. Artificial Neural Networks, Parzen Windows, Hidden Markov Random Fields etc. are required to training data. Unsupervised techniques do not require training data. The Fuzzy Clustering Techniques are unsupervised techniques and they calculate the clusters and cluster centers with considering the mean and distance values (Duda Richard O.,2001).

In this thesis, the goal is, segmentation of the MS lesion areas from healty brain tissues using pattern recognition techniques from T1, T2 and PD weighted MRI images. These techniques are:

- Fuzzy C-Means clustering
- K-Means clustering
- K-Medoid clustering
- Principle component analysis
- Independent component analysis
- Artificial neural networks

Comparison of the supervised techniques (Multilayer Perceptron Backpropagation Neural Networks) and unsupervised techniques (Fuzzy C-Means, K-Means, K-Medoid, Principle Component Analysis and Independent Component Analysis) are evaluated. The results are compared according to their sensitivity, accuracy and time and from the all conserened tecniques, optimum one is suggested.

In the first chapter, our motivation is given. In the second chapter; Methodology of magnetic resonance imaging is considered; The third chapter, classification of the methods for brain tissues specially MS lessions by using MRI images summarizes; The forth chapter, our application methods and their analysis results depending on several criteria are given. Finally in the fifth chapter, evaluation of the results and our suggestions are presented.

CHAPTER TWO
MAGNETIC RESONANCE IMAGING

2.1 Magnetic Resonance Imaging

MRI (Magnetic Resonance Imaging) is a nonionizing technique with full three dimensional capabilities. The major uses of MRI are in the areas of assessing brain disase, spinal disorders, angiography, cardiac function, and musculoskeletal damage (Hornak J.P., 2007).

The patient is placed inside a strong magnetic field about 1.5 T. (10000 times strong then earth's magnetic field).

In general MRI is based on a property of the nucleus called magnetic spin moment. Any charged spinning particle creates an electromagnetic field, like a bar magnet with the magnetic field pointing from the South Pole to the North Pole. Only charged nuclei have this property. The single proton of hydrogen has two possible energy states, one resulting from each direction of spin (Figure 2.1).

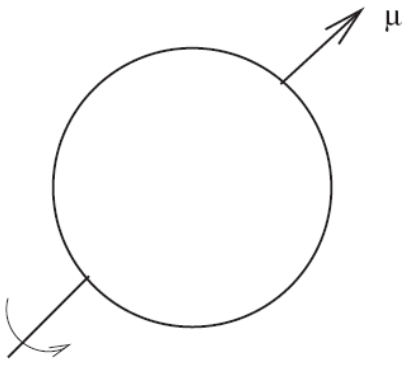


Figure 2.1 A single proton with spin, produces a magnetic moment, μ..
(Mbah, Henry Osita, 2006)

In order to obtain an MRI signal, transitions must be induced between the protons in the parallel and antiparallel energy levels. To produce a signal, magnetic energy must be supplied to the protons at larmor frequency in order to stimulate transitions.

$$w = \gamma B_0 \qquad \text{(Eq. 2.1)}$$

w = larmor frequency

γ = gyromagnetic ratio of the nucleus

B_0 = magnetic field

Other nuclei with different number of protons, will have more energy states. Only elements with unpaired number of protons can be used in MRI, if there were an even number of protons, the spins would cancel each other out. Hydrogen is the element being used, because of its high abundance in all human tissues.

All protons in hydrogen will create their own magnetic field, called magnetic dipole moment. If there is no external magnetic field, the magnetic dipole moment will cancel out because of the high number of nuclei, and there will be no net magnetic field. If we then impose an external field B_0 in the z-direction, the protons will align towards the north or to the south. But around one in a million extra spins will point north, creating a net magnetisation in the direction of B_0. Actually, each proton will neither point directly towards north, nor directly towards south, but will begin to precess around the z-axis, see Figure 2.2.

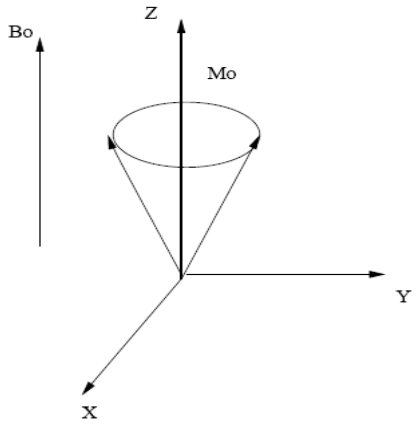


Figure 2.2 Precession of the magnetic dipole moment around the z-axis when exposed to a static magnetic field B_0. (Mbah, Henry Osita,2006).

However, the protons are not in phase, and this causes a net magnetisation in the direction of the external field B_0. The frequency at which the proton precesses around the axis of the external magnetic field is given by the Larmor equation.

Consider a hydrogen proton being placed in a strong static magnetic field B_0 along the z-axis. If upon the static magnetic field we transmit a RF (Radio Frequency) pulse along the x-axis perpendicular to the z-axis, the proton will start to precess around the axis of the RF field. Consequently the proton will begin to precess about the x-axis. This precessional frequency will be:

$$W_o = \gamma \mid B_1 \mid \qquad \text{(Eq. 2.2)}$$

B_1 = magnetic field of the RF pulse.

B_1 (e.g. 50mT) is oscillating and weak compared to B_0 (1.5T). These two precessional frequencies result in a spiral motion of the net magnetisation vector from the z axis into the xy plane.

If the frequency of the RF is equal to the Larmor frequency W_o of B_0, then resonance occurs, and the RF pulse will add energy to the proton. The magnetization vector will flip into the x-y plane, and the protons will then be in phase. This causes transversal magnetization. In MRI we transmit such a RF with the correct frequency causing resonance, and we can decide the flip angle from the duration and the strength of the pulse.

When the RF pulse is turned off, the protons will proceed back into their state of equilibrium under the influence of the static magnetic field B_0. This process creates an electromagnetic pulse from the protons, and the strength of this pulse is dependent on the concentration of protons in that specific area where the signal is generated. RF coils which is usually placed directly around, or next to, the tissue to be imaged delivers this energy . RF coil is designed for stored as much of the magnetic enery as posibble in the near field region. The most efficient coil design is based on resonance electric circuits which has a resonant frequency w_r that stored in the coil

is maximum.

$$w_r = \frac{1}{\sqrt{LC}} \qquad \text{(Eq. 2.3)}$$

L = inductance of the RF coil

C = capacitance

This electromagnetic pulse is measured by the coil surrounding the patient, and the strength of this pulse is proportional to the proton density, hence we can create an image which depends on the proton density. To ensure spatial information from the received signal and locate it in space, gradient coils are used to create small perturbations of the static magnetic field. Three orthogonal coils are used to achieve this. Because of these gradients of the static field, each RF frequency will have a unique volume of response, and the signal is unique with regard to location.

RF coils can be classified in according to their usages and geometries. In according to their geometry, surface coil, quadrature surface coil, soleneid and birdcage coil are some of them. RF coils are also classified as transmit, receive, transmit-receive and multicoil arrays. A multicoil array typically receives only and they are used to increase sensitivity.

The readout MR signal is a mix of RF waves with different amplitudes, frequencies and phases, containing spatial information. This signal is digitalized and raw data are written into a data matrix called K-space. K-space data are equivalent to a Fourier plane. To go from a k-space data to an image requires using a 2D inverse Fourier Transform.

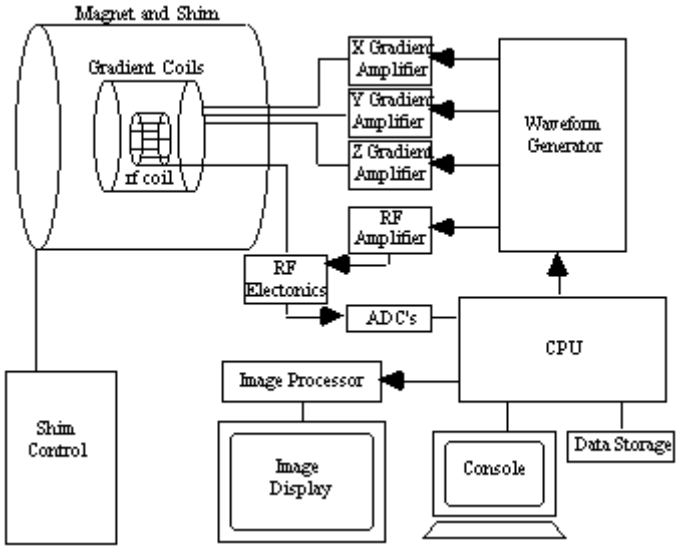
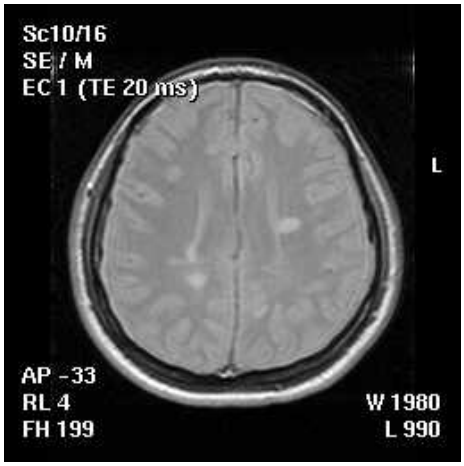


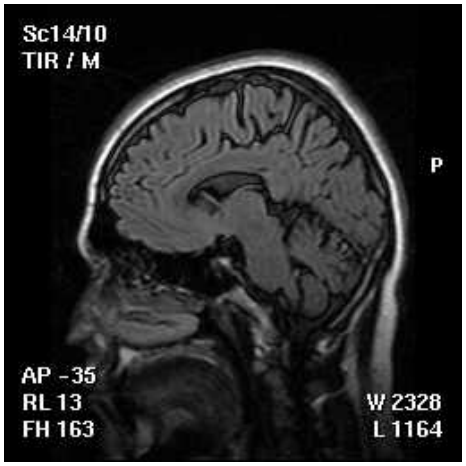
Figure 2.3 Schematic diagram of a MRI scanner. (http://users. fmrib.ox.ac.uk/~stuart/thesis/chapter_2/section2_6.html#3)

Frequency-encoding and phase-encoding are done so that data is spatially encoded by differences in frequency and phase, amenable to analysis by Fourier transform. In k-space, fx-coordinates (horizontal spatial frequencies) and fy-coordinates (vertical spatial frequencies) of the Fourier plane are replaced by kx and ky-coordinates.

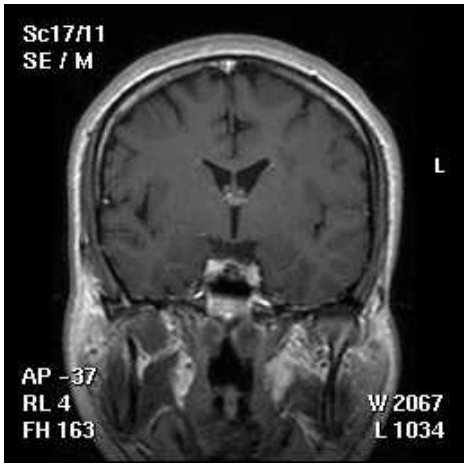
Clinical MRI studies acquire a series of slices through the anatomical area of interest, each slice has tickness. Slice selection directions are important and accomplishes by RF puls application to the magnetic field gradients. The choice of slice select direction of the images can be obtained by coronal (y), axial (z) and sagittal (x).



(a)



(b)

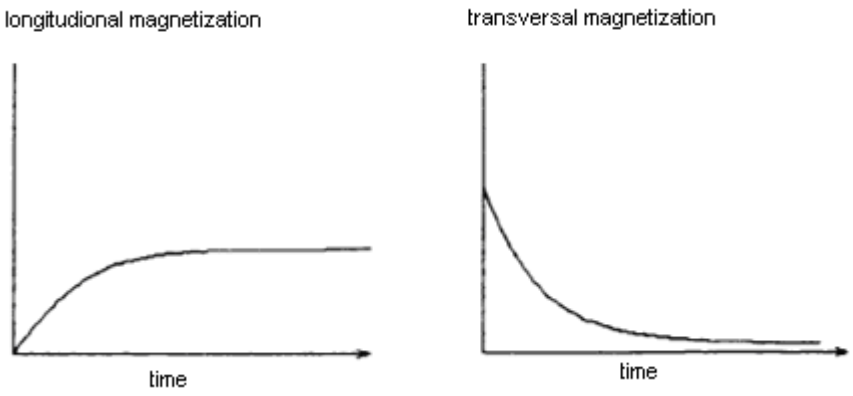


(c)

Figure 2.4 (a) Axial image. (b) Sagittal image. (c) Coronal image (MRI Images were taken from GATA.)

With the emission of MR signals, the spin system returns to the low energy state usually by two spin relaxation mechanisms: longitudinal relaxation and transverse relaxation. In longitudinal relaxation, when the RF pulse is switched off, the longitudinal magnetization increases. The time that takes for the longitudional magnetization to go back to its original value is described by the longitudinal relaxation time, called T1. At the same time transfer magnetization decreases and

disappeares. The time constant describing how fast transversal magnetization vanishes is called the transverse magnetization relaxation time. Figure 2.5 shows the longitudinal and transverse magnetization curves versus time for T1 and T2 weighted images.

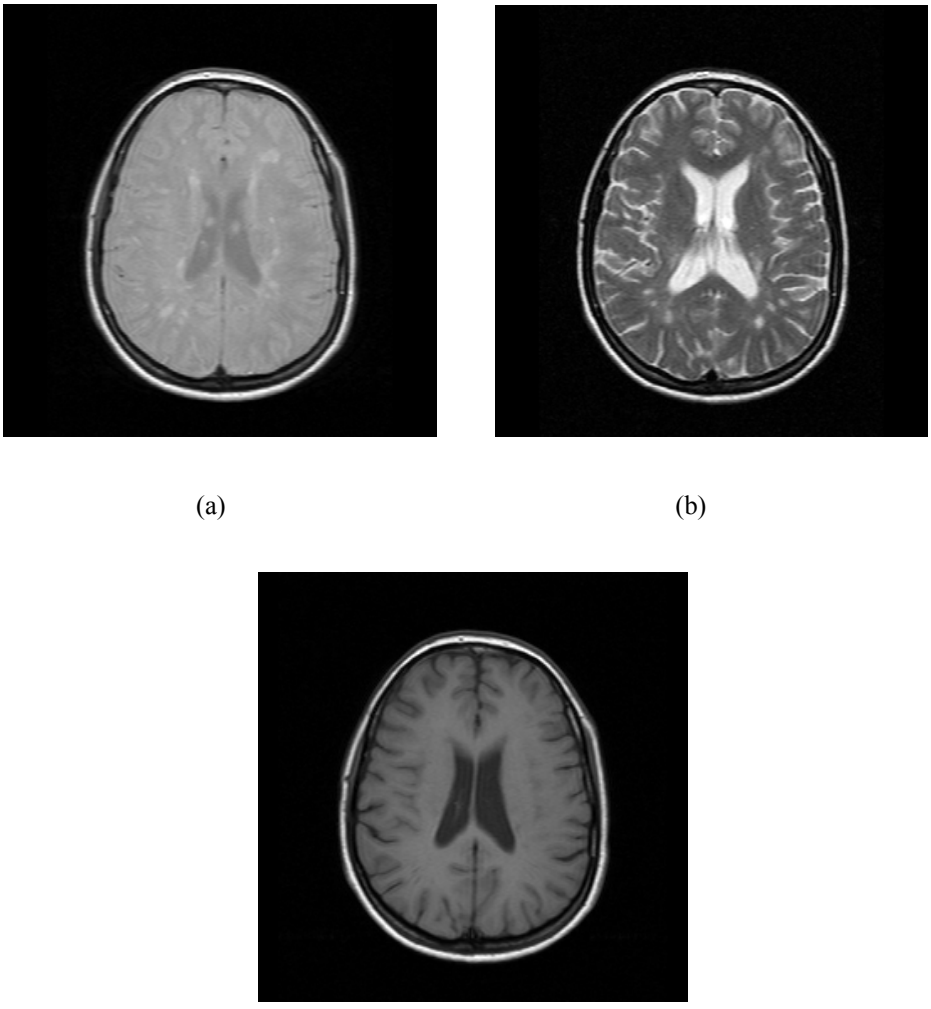


(a) (b)

Figure 2.5 longitudinal and transverse magnetization curves versus time for T1 and T2 weighted images (Pu Ma, 2001) .

Appropriately time puls sequences can give three types of images. T1 weighted, Proton Density and T2 weighted. The spin echo puls sequences is the most commonly used pulse sequences and the dual echo spin pulse sequences can be used to obtain both T2 and PD weighted images simultaneously. All spin echo sequences include a slice selective 90 ^{o} puls followed by one or more 180 ^{o} pulses. The 180^{o} pulses here acts as a wall from which the protons bounce back like echoes. Thus, the resulting strong signal is called an echo, or spin echo. The repetition time TR (a period between two puls sequences) and the echo time TE (a period between the start of the RF puls and echo) determine how the resulting image is weighted. A short TR and short TE will give a T1 weighted image. A long TR and short TE will give a proton density image. A long TR and long TE will give a T2 weighted image.

T1 weighted images have very high contrast between cerebrospinal fluid (CSF) and other brain tissues. However lesions are not easly identifiable. In PD images the CSF is dark with high contrast to both gray matter and lesions. The contrast between other tissues of interest such as between gray matter or white matter or between lesions and gray matter is not very high. In T2 weighted images both CSF and lesions appear as bright regions and the contrast between the brain parenchyma and the CSF is high. Figure 2.6 shows the T1, T2 and PD MRI images.



(a) (b)

(c)

Figure 2.6 (a) PD (b) T2 (c) T1 weighted images. (MRI Images were taken from GATA.)

CHAPTER THREE
IMAGE CLUSTERING METHODS

Pattern recognition techniques can be classified into two broad cathegories: unsupervised and supervised techniques. An unsupervised technique doesn't use a prior class identifiers, whereas a supervised technique uses a data set with known data classifications (Yen J., 2000). These two types of techniques are complementary. For example, unsupervised clustering can be used to produce classification information needed by a supervised pattern recognition technique. In the area of pattern recognition and image processing, unsupervised clustering is often used to perform the task of segmenting the images. This is because image segmentation can be viewed as kind of data clustering problem where each data point describe a set of image features associated to each pixel.

3.1 Unsupervised Learning and Clustering

Unsupervised clustering is motivated by the need to find interesting patterns or groupings in a given set of data.

Conventional clustering algorithms find a hard partition of given dataset based on certain criteria that evaluate the goodness of a partition. Hard partition means that each data belong to exactly one cluster of the partition.

Let X be a set of data and x_i be an element of X. A partition $P=\{C_1,C_2,\ldots..C_L\}$ of X is "hard" if and only if;

$\forall x_i \in X \quad \exists C_j \in P$ such that $x_i \in C_j$

This condition assures that the partition covers all data points in X. However, if;

$\forall x_i \in X \quad x_i \in C_j \Rightarrow x_i \notin C_i$ where $k \neq j$, $C_k, C_j \in P$ then all clusters in the partition are mutually exclusive.

In many real world clustering problems some data points partially belong to multiple clusters rather then a single cluster exclusively. For example, a pixel in a magnetic resonance image may corresponde to mixture of a different types of tissues.

A soft clustering algorithm finds a soft partition of a given data setbased on certain criteria. In soft partition, a data can partially belong to multiple clusters.

$\forall x_i \in X, \quad \forall C_j \in P \qquad 0 \leq \mu_{C_j}(x_i) \leq 1$

$\forall x_i \in X, \quad \exists C_j \in P$ such that $\mu_{C_j}(x_i) > 0$, where $\mu_{C_j}(x_i)$ denotes the degree to which x_i belongs to cluster C_j.

A type of soft clustering of special interest is one that ensures the membership degree of a point x in all clusters adding up to one,

$$\sum_j \mu_{C_j}(x_i) \qquad \forall x_i \in X \qquad \text{(Eq 3.1)}$$

A soft partition that satisfies this additional condition is called a constrained soft partition (Cuevas J. Erik, 2004)

3.2 Clustering Algorithms

3.2.1 K-means and K-medoid algorithms

The K-means and K-medoid hard partitioning methods are simple and popular methods. From an Nxn dimensional data set K-means and K-medoid algorithms allocates each data point to one of c clusters to minimize the within cluster sum of squares:

$$\sum_{i=1}^{c} \sum_{k \in A_i} \left\| x_k - v_i \right\|^2 \qquad \text{(Eq 3.2)}$$

A_i is a set of data points in the ith cluster

v_i mean for that points over cluster i (cluster prototypes)

$$v_i = \frac{\sum_{k=1}^{N_i} x_k}{N_i}, x_k \in A_i \qquad \text{(Eq 3.3)}$$

N_i number of objects in A_i.

In K-medoid clustering the cluster centers are the nearest objects to the mean of data in one cluster (Balasko B.,2006).

3.2.2 Fuzzy C-means algorithm

The Fuzzy C-means clustering algorithm is based on the minimization of an objective function called C-means functional (Bezdek J. C., 1975).

$$J(X;U,V) = \sum_{i=1}^{c}\sum_{k=1}^{N} (\mu_{ik})^m \left\| x_k - v_i \right\|_A^2 \qquad \text{(Eq 3.4)}$$

Where

$V = [v_1, v_2, ...v_c], v_i \in R^n$,

V is a vector of cluster centers (prototypes).

Squared inner-product distance norm:

$$D_{ikA}{}^2 = \left\| x_k - v_i \right\|_A^2 = (x_k - v_i)^T A (x_k - v_i) \qquad \text{(Eq 3.5)}$$

From equation 3.4 total variance of x_k from v_i can be seen. The minimization of the c-means functional represents a non-linear optimization problem that can be solved by using a variety of available methods, ranging from grouped coordinate minimization, over simulated annealing to genetic algorithms. The most popular method, however, is a picard iteration through the first order conditions for stationary points of equation 3.4, known as fuzzy c-means (FCM) algorithm.

J by means of Lagrange multipliers:

$$J(X;U,V) = \sum_{i=1}^{c}\sum_{k=1}^{N} (\mu_{ik})^m D_{ikA}{}^2 + \sum_{k=1}^{N} \lambda_k \left(\sum_{i=1}^{c} \mu_{ik} - 1\right) \qquad \text{(Eq 3.6)}$$

If $D_{ikA}{}^2 > 0, \forall i,k$ and $m > 1$, then $(U,V) \in M_{fc} \times R^{n \times c}$ may minimize Equation 3.4 only if,

$$\mu_{ik} = \frac{1}{\sum_{j=1}^{c} (D_{ikA} / D_{jkA})^{2/(m-1)}}, \; 1 \le i \le c, 1 \le k \le N \qquad \text{(Eq 3.7)}$$

$$v_i = \frac{\sum_{k=1}^{N} \mu_{ik}^m x_k}{\sum_{k=1}^{N} \mu_{ik}^m}, \quad 1 \le i \le c \qquad \text{(Eq 3.8)}$$

This equation gives v_i as the weighted mean of the data items that belong to a cluster, where the weights are the membership degrees. The FCM algorithm is a simple iteration through equation 3.7 and 3.8.

The FCM algorithm computes with the standard Euclidian distance norm, which induces hyperspherical clusters. Hence it can only detect clusters with the same shape and oriantation, because the common choice of norm inducing matrix is A=I or it can be chosen as an nxn diagonal matrix that accounts for different variances in the directions of the coordinate axes of X.

$$A_D = \begin{bmatrix} (1/\sigma_1)^2 & 0 & \dots & 0 \\ 0 & (1/\sigma_1)^2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & (1/\sigma_1)^2 \end{bmatrix} \qquad \text{(Eq 3.9)}$$

A can be defined as the inverse of the nxn covariance matrix.

$$A = F^{-1} \qquad \text{(Eq 3.10)}$$

$$F = \frac{1}{N} \sum_{k=1}^{N} (x_k - \bar{x})(x_k - \bar{x})^T \qquad \text{(Eq 3.11)}$$

Here $\bar{x}$ denotes the sample mean of the data. In this case, A induces the mahalanobis norm on R^n (Balasko B.,2006).

3.3 Component Analysis

Component analysis is an unsupervised approach to finding the right features from the data. Principle component analysis projects d-dimensional data onto a lower dimensional subspace in a way that is optimal in a sum squared error sense. Non linear component analysis typically implemented by neural techniques is a non linear generalization of principle component analysis. In independent component anaysis those directions in feature space that show the independence of signals. This method is particularly helpful for segmenting signals from multiple sources (Duda Richard O., 2001).

3.3.1 Principle Component Analysis (PCA)

Principle component analysis (PCA) involves a mathematical procedure that transforms a number of possibly corraleted variables into a smaller number of uncorraleted variables called principle components. The first principle component accounts for as much of the variability in the data as possible, and each succeeding component accounts for as much of the remaining variability as possible. The main objectives of PCA are (Balasko B.,2006):

- Identify new meaningful underlying variables
- Discover or reduce the dimensionality of the data set.

If we assume that representing all of the vectors in a set of n dimensional samples $x_1,\ldots,x_n$ by a single vector x_0. To find a vector x_0 such that the sum of the squared distances between x_0 and the various x_k is as small as possible. The squared error criterion function is (Duda Richard O., 2001):

$$J_0(x_0) = \sum_{k=1}^{n} \left\| x_0 - x_k \right\|^2 \qquad \text{(Eq 3.12)}$$

That computes the x_0 that minimizes J_0.

If x_0=m, where m is the sample mean:

$$m = \frac{1}{n}\sum_{k=1}^{n} x_k \qquad \text{(Eq 3.13)}$$

$$J_0(x_0) = \sum_{k=1}^{n} \left\| (x_0 - m) - (x_k - m) \right\|^2 \qquad \text{(Eq 3.14)}$$

$$= \sum_{k=1}^{n} \left\| x_0 - m \right\|^2 - 2\sum_{k=1}^{n} (x_0 - m)^t (x_k - m) + \sum_{k=1}^{n} \left\| x_k - m \right\|^2 \qquad \text{(Eq 3.15)}$$

$$= \sum_{k=1}^{n} \left\| x_0 - m \right\|^2 - 2(x_0 - m)^t \sum_{k=1}^{n} (x_k - m) + \sum_{k=1}^{n} \left\| x_k - m \right\|^2 \qquad \text{(Eq 3.16)}$$

$$= \sum_{k=1}^{n} \left\| x_0 - m \right\|^2 + \sum_{k=1}^{n} \left\| x_k - m \right\|^2 \qquad \text{(Eq 3.17)}$$

The second part of the equation 3.17 is independent of x_0. So that the choice of x_0=m minimizes the expression.

The sample mean is a zero dimensional representation of the data set. For another representation a unit vector can be used.

e is the unit vector in the direction of the line. The equation of line is:

$$x = m + \alpha e \qquad \text{(Eq 3.18)}$$

α is corresponds to the distance of any point x from the mean m. The representation of x_k by $m + a_k e$ than optimal set of coefficients can be calculated by minimizing the squarred error criterion function:

$$J_1(a_1, ..., a_n, e) = \sum_{k=1}^{n} \left\| (m + a_k e) - x_k \right\|^2 = \sum_{k=1}^{n} \left\| a_k e - (x_k - m) \right\|^2 \qquad \text{(Eq 3.19)}$$

$$= \sum_{k=1}^{n} a_k^2 \left\| e \right\|^2 - 2\sum_{k=1}^{n} a_k e^t (x_k - m) + \sum_{k=1}^{n} \left\| x_k - m \right\|^2 \qquad \text{(Eq 3.20)}$$

If $\left\| e \right\|$=1 then a_k can be obtain as:

$$a_k = e^t (x_k - m) \qquad \text{(Eq 3.21)}$$

The meaning of the result is the projecting of the x_k onto the line in the direction of e.To finding best direction of e, the scatter matrix (S) can be used:

$$S = \sum_{k=1}^{n} (x_k - m)(x_k - m)^t \qquad \text{(Eq 3.22)}$$

If S is used in the squared error criteria:

$$J_1(e) = \sum_{k=1}^{n} a_k^2 - 2\sum_{k=1}^{n} a_k^2 + \sum_{k=1}^{n} \left\| x_k - m \right\|^2 \quad \text{(Eq 3.23)}$$

$$= -\sum_{k=1}^{n} [e^t (x_k - m)]^2 + \sum_{k=1}^{n} \left\| x_k - m \right\|^2 \quad \text{(Eq 3.24)}$$

$$= -\sum_{k=1}^{n} e^t (x_k - m)(x_k - m)^t e + \sum_{k=1}^{n} \left\| x_k - m \right\|^2 \quad \text{(Eq 3.25)}$$

$$= -e^t S e + \sum_{k=1}^{n} \left\| x_k - m \right\|^2 \quad \text{(Eq 3.26)}$$

The vector e, that minimizes J_1 also maximizes $e^t S e$.

Let λ be the undetermined multiplier. The differantiate with respect to e:

$$u = e^t S e - \lambda (e^t e - 1) \quad \text{(Eq 3.27)}$$

$$\frac{\partial u}{\partial e} = 2Se - 2\lambda e \quad \text{(Eq 3.28)}$$

Setting this gradient vector equal to zero, e must be an eigenvector of the scatter matrix:

$$Se = \lambda e \quad \text{(Eq 3.29)}$$

To find the best dimensional projection of the data it should be projected onto a line through the sample mean in the direction of the eigenvector of the scatter matrix having the largest eigenvalue. This result can be readily extended from a one dimensioanl projection to a $d^/$ dimensional projection :

$$x = m + \sum_{i=1}^{d^/} a_i e_i \quad \text{(Eq 3.30)}$$

Where $d^/ \le d$. The criterion function is:

$$J_{d^/} = \sum_{k=1}^{n} \left\| \left(m + \sum_{i=1}^{d^/} a_{ki} e_i \right) - x_k \right\|^2 \quad \text{(Eq 3.31)}$$

Minimized when the vectors $e_{1,}....e_{d^/}$ are the $d^/$ eigenvectors of the scatter matrix having the largest eigenvalues. Because the scatter matrix is real and symmetric, these eigenvectors are orthogonal. They form a naturel set of basis vectors for

representing any feature vector x. The coefficients a_i are the component of x and they are called to principle components (Duda Richard O., 2001).

3.3.2 Nonlinear Component Analysis (NLCA)

Principle component analysis yields a k-dimensional linear subspace of feture space that best represents the full data according to a minimum square error criterion(Duda Richard O., 2001). If the data represent complicated interactions of features, then the linear subspace may be a poor representation and nonlinear components may be needed. This can be obtained with a neural network approach which has d input and d output layers and k<d middle layer. In this middle layer nonlinear components can be revealed.

3.3.3 Independent Component Analysis (ICA)

Independent component analysis finds the most independent mapping directions from each other.

Suppose there are d independent scalar source signals $x_i(t)$ for i=1,2,…,d where time index t is 1<t<T.

The multivariate density function is (Duda Richard O., 2001):

$$p[x(t)] = \prod_{i=1}^{d} p[x_i(t)] \qquad \text{(Eq 3.32)}$$

If there is k dimensional data, and A is a kxd matrix then:

$$s(t) = Ax(t) \qquad \text{(Eq 3.33)}$$

The goal of ICA is to extract d components in s that are independent.

If k=d is choosen, the distrubution in the outputs is related to the distribution where J is the Jacobian matrix:

$$p_y(y) = \frac{p_s(s)}{|J|} \qquad \text{(Eq 3.34)}$$

$$J = \begin{pmatrix} \frac{\partial y_1}{\partial s_1} & \cdots & \frac{\partial y_d}{\partial s_1} \\ \vdots & \vdots & \vdots \\ \frac{\partial y_1}{\partial s_d} & \cdots & \frac{\partial y_d}{\partial s_d} \end{pmatrix} \quad \text{(Eq 3.35)}$$

$$|J| = \left| |W| \prod_{i=1}^{d} \frac{\partial y_i}{\partial s_i} \right| \quad \text{(Eq 3.36)}$$

The final stage as a linear transform of the source signals:

$$y = f[W_s + w_0] \quad \text{(Eq 3.37)}$$

w_0 is the bias vector and $f[.]$ is typically choosen as sigmoid. The central goal in ICA is to find the parameters W and w_0 so as to make the outputs y_i as independent from one to another as posibble (Duda Richard O., 2001) .

3.4 Supervised Learning

3.4.1 Multilayer Neural Networks

Neurocomputing is concerned with processing information. Unlike its programmed computing counterpart, a neurocomputing approach to information processing first involves a learning process within an artificial neural network architecture that adaptively responds to inputs according to a learning rule. After a neural network has learned what it needs to know, the trained network can be used to perform certain tasks depending on the particuliar application. Neural networks have the ability to learn from their environment and to adapt to it in an interactive manner similar to their biological counterparts. Indeed, this is an exiciting prospect because of the vast possibilities that exist for performing certain functions with artificial neural networks that can emulate the comparable biological function (Ham Fredric M., 2001).

3.4.1.1 A Neuron Model

An artificial neuron can be referred to as a processing element or a threshold logic unit. A neuron is an information processing unit that roughly resembles its biological counterpart. There are four basic components of the model.

The continous valued input to the synapsis is a vector a vector signal $x \in R^{n\times 1}$, with the individual vector components given as x_j, for j=1,2,…,n, that is $x = [x_1, x_2, ..., x_n]^T$. Therefore, each vector component x_j is the input to the jth synapse and connected the neuron q through a synaptic weight w_{qj}, that is x_j is multiplied by the syneptic weight w_{qj}. The summing device acts as add all the signals that are broadcast in to the adder. Output of the adder u_q, is a linear combination of inputs to the synepsis. The activation function $f(.)$ serves to limiting the amplitude of the neuron output y_q when $f(.)$ is a nonlinear function. The threshold θ_q is usually externally applied and lowers the cummulative input to the activation function. Therefore, θ_q is subtracted from the output of the linear combiner u_q before the activation is applied. Figure 3.1 shows a simple neuron model.

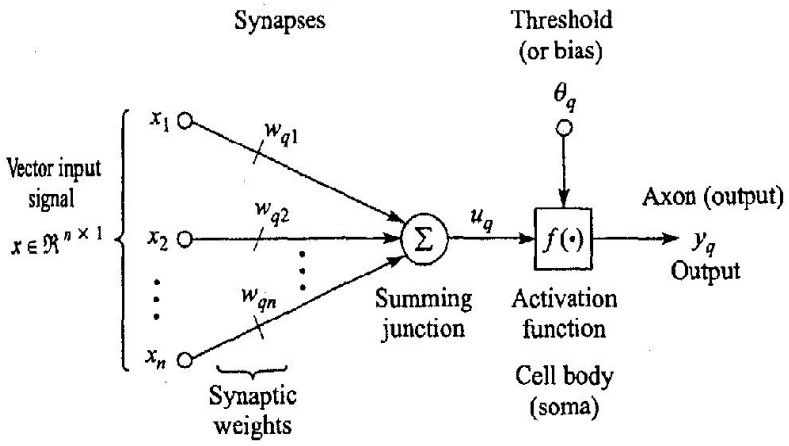


Figure 3.1 Nonlinear model of an artificial neuron. (Frederic M. Ham, Ivica Costanic "Principles of Neurocomputing for Science and Engineering" p.25)

Mathematically the operation of the artificial neuron can be described as :

$$u_q = \sum_{j=1}^{n} w_{qj} x_j = w_q^T x = x^T w_q \qquad \text{(Eq 3.38)}$$

Where $x = [x_1, x_2, ..., x_n]$ and $w_q = [w_{q1}, w_{q2}, ..., w_{qn}]^T \in R^{n\times 1}$.The output activation function is :

$$y_q = f(v_q) = f(u_q - \theta_q) \qquad \text{(Eq 3.39)}$$

$$y_q = f\left(\sum_{j=1}^{n} w_{qj} x_j - \theta_q \right) \qquad \text{(Eq 3.40)}$$

3.4.1.2 Backpropagation Learning Algorithm (MLP NN)

Backpropagation is the most widely used learning process in neural networks. Figure 3.2 shows an example of the multilayer perceptron neural network architecture. The extension of the derivation to the general case when the network has more than two hidden layers is straightforward (Ham Fredric M., 2001).

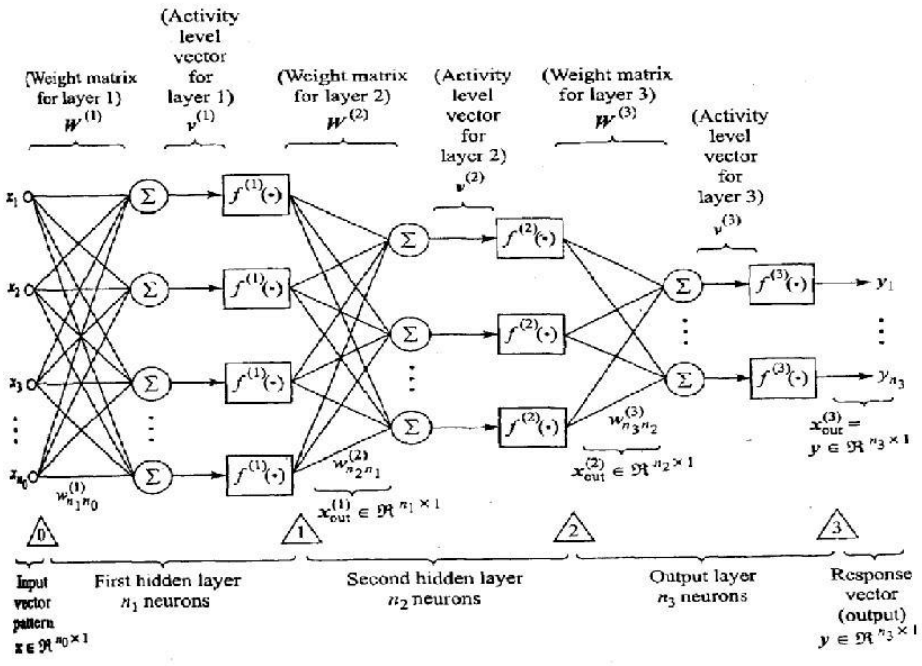


Figure 3.2 An example of a three layer feedforward MLP NN architecture.
(Frederic M. Ham, Ivica Costanic "Principles of Neurocomputing for Science and Engineering" p.107)

The standard backpropagation algorithm for training of the MLP NN is based on the steepest descent gradient approach applied to the minimization of an energy function representing the instantaneous error (Ham Fredric M., 2001).

$$E_q = \frac{1}{2}\left(d_q - x_{out}^{(3)}\right)^T \left(d_q - x_{out}^{(3)}\right) = \frac{1}{2}\sum_{h=1}^{n_3}\left(d_{qh} - x_{out,h}^{(3)}\right)^2 \qquad \text{(Eq 3.41)}$$

d_q is the desired network output for the qth input pattern and $x_{out}^{(3)} = y_q$ is the actual output of the MLP network shown in figure 2.2. In general the method which has derived at Eq 2.48 is called the online method and it has minimum memory storage requeriments.

Using the steepest descent gradient approach, the learning rule for a network weight in any one of the network layer is:

$$\Delta w_{ji}^{(s)} = -\mu^s \frac{\partial Eq}{\partial w_{ji}^{(s)}} \qquad \text{(Eq 3.42)}$$

S is the appropriate network layer and $\mu > 0$ is the corresponding learning rate parameter.

Derivations of the seperate learning rules for weights in the output and of the MLP NN is:

$$w_{ji}^{(s)}(k+1) = w_{ji}^{(s)}(k) + \mu^s \delta^{(s)} x_{out,i}^{(s)} \qquad \text{(Eq 3.43)}$$

$$\delta_j^{(s)} = (d_{qh} - x_{out,i}^{(s)}) g(v_j^{(s)}) \qquad \text{(Eq 3.44)}$$

For hidden layers:

$$\delta_j^{(s)} = \left(\sum_{h=1}^{n_{s+1}} \delta_h^{(s+1)} w_{hj}^{(s+1)}\right) g(v_j^{(s)}) \qquad \text{(Eq 3.45)}$$

Where g(.) is the first dervative of the f(.).

CHAPTER FOUR
MS SEGMENTATION

4.1 The Data

For applications of the segmentation techniques, quantitative (numerical) data is considered. The data used in the applications is taken from the McGill University Simulated Brain Data Base (BrainWeb: Simulated Brain Database, http://www.bic.mni.mcgill.ca/brainweb/), (April 2007). These data sets are T1, T2, PD weighted and for supervised techniques, target MS lesions labeled MRI images with 181x217x181 sized. Each MRI image set has 181 x217 images and 181 slices. According to the application techniques, images are reshaped to new matrices as Nxn size. N=181x217 and n is the observation number.

Figure 4.1 shows the 105. slices T1,T2,PD weighted bands and target (MS lesions) images. Also Figure 4.2 shows the 100. slices T1,T2,PD weighted bands and target (MS lesions) images.

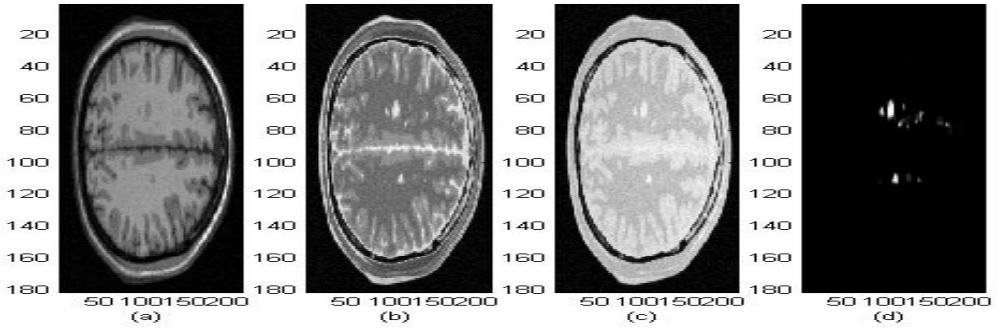


Figure 4.1 (a) T1 weighted (B) T2 weighted (c) PD (d) Target MRI images

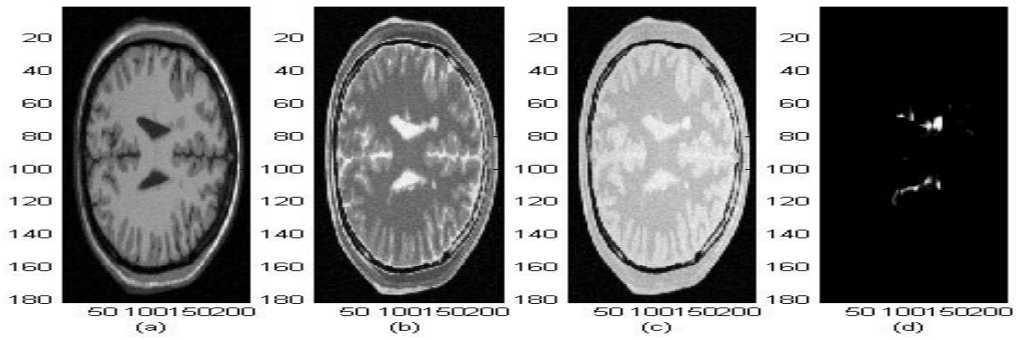


Figure 4.2 (a) T1 weighted (B) T2 weighted (c) PD (d) Target MRI images

4.2 Segmentation Methods

Segmentation methods can be classified as supervised and unsupervised techniques. Unsupervised methods are known as clustering techniques. Supervised methods need labeled data for training. The techniques, which were used in this study, are given below:

Unsupervised Techniques:

- Fuzzy C Means Clustering
- K-Means Clustering
- K-Medoid Clustering
- Principle Component Analysis
- Independent Component Analysis

Supervised Technique:

- Feedforward Multilayer Perceptron Neural Network

Figure 4.3 shows these segmentation algorithms.

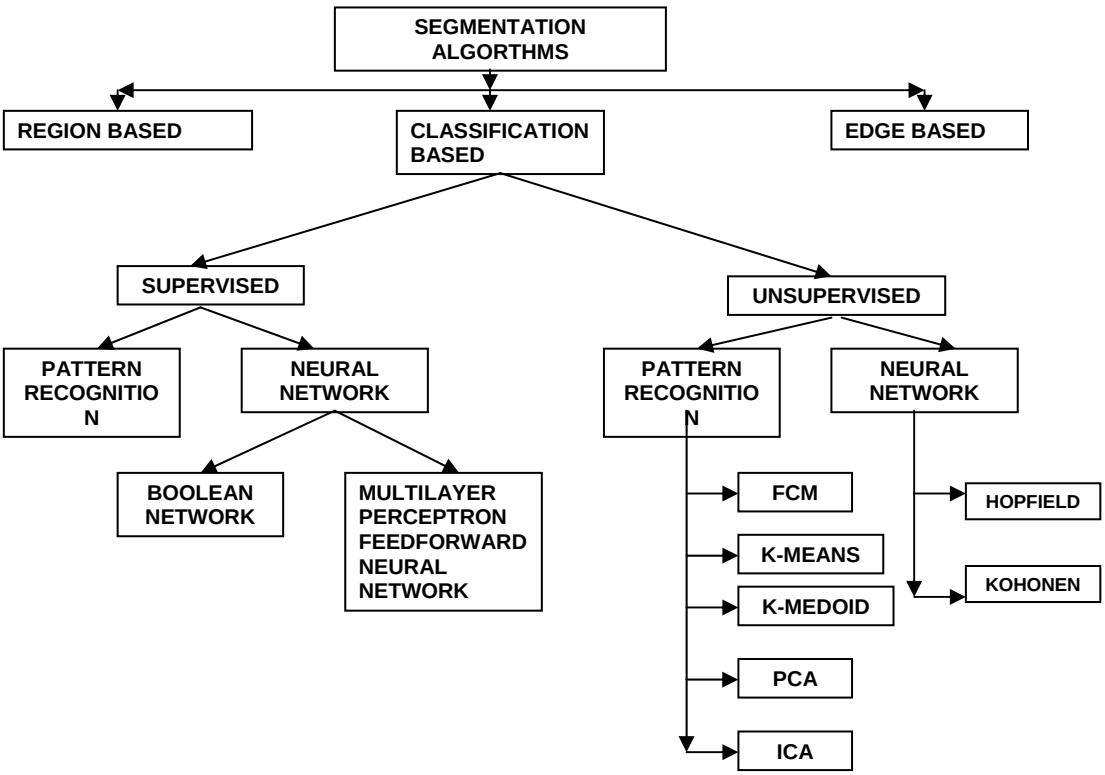


Figure 4.3 Segmentation algorithms

4.3 Input Data

In order to obtain clustering analysis, first of all, new input data must be composed to application of the segmentation process (Özkurt A., 2007). These input data are related to the T1,T2 and PD images of the specific MRI image with different nonlinear equations, like square root or logaritmic functions(Ren, H.,2000). These band composition techniques are given below (Özkurt A., 2007):

$a) B_i^- = \{255 - B_i\}_{i=1}^3$

$b) sqrt\{B_i\}_{i=1}^3$

$c) sqrt\{B_i B_j\}_{i,j=1}^3, i \neq j$

$d) \{B_i^2\}_{i=1}^3$

$e) \{B_i B_j\}_{i,j=1}^3, i \neq j$

$f) \{B_i B^-{}_j\}_{i,j=1}^3, i \neq j$

Where B_1 is the T1 weighted reshaped MRI image, B_2 is the T2 weighted reshaped MRI image and B_3 is the PD weighted reshaped MRI image.

4.4 Segmentation With Clustering Techniques

4.4.1 Fuzzy C Means Clustering

The Fuzzy C-Means Clustering algorithm uses the minimization of the Fuzzy C-Means functional. The function calculates the standard Euclidian distance norm, the norm inducing matix is an nxn identity matrix. The result of the partition is collected in structure arrays.

If c is the cluster number, first of all the cluster means are calculated with:

$$v_i^{(l)} = \frac{\sum_{k=1}^{N} (\mu_{ik}^{(l-1)})^m x_k}{\sum_{k=1}^{N} (\mu_{i,k}^{(l-1)})^m}, 1 < i < c \qquad \text{(Eq 4.1)}$$

Where $l = 1, 2, ..$

The distences between pixels and cluster means are:

$$D_{ikA}^2 = (x_k - v_i)^T A(x_k - v_i), 1 \le i \le c, 1 \le k \le N \quad \text{(Eq 4.2)}$$

Where A is the norm-inducing matrix.

The partition matrix is:

$$\mu_{i,k}^{(l)} = \frac{1}{\sum_{j=1}^{c} \left(\frac{D_{ikA}}{D_{jkA}} \right)^{\frac{2}{(m-1)}}} \quad \text{(Eq 4.3)}$$

Algorithm for the fuzzy c-means clustering is shown in the figure 4.4.

Fuzzy c-means clustering is accomplished using MATLAB fuzzy clustering toolbox (Balasko B., 2006). In the application stage, first of all, 100. slice has been used for clustering. Input data for the clustering is non-linear combination of the T1-T2-PD images (Özkurt A., 2007). Some of the results are shown in the figure 4.5.

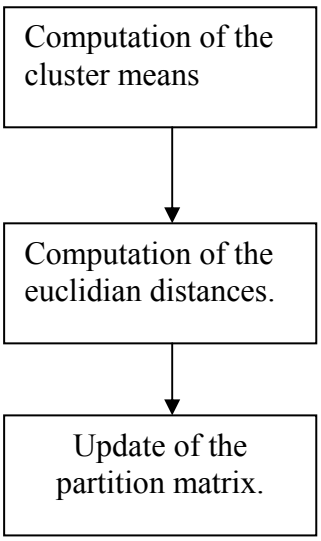


Figure 4.4 Algorithm for the FCM.

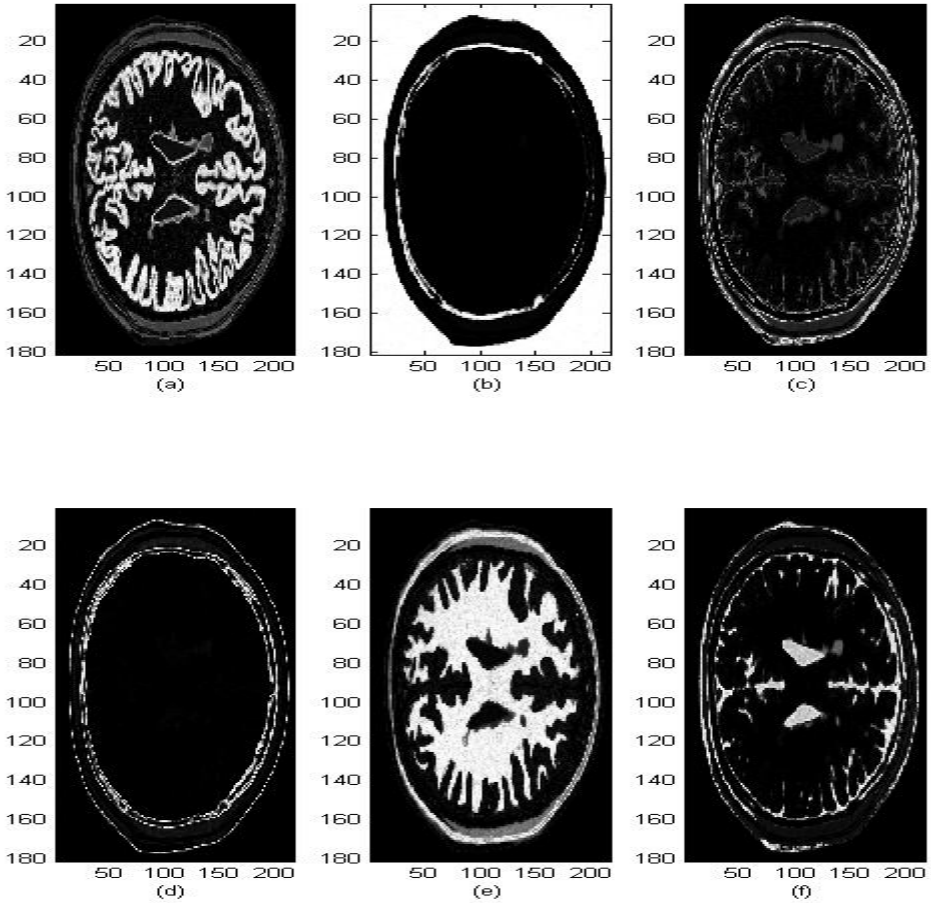


Figure 4.5 Result cluster classes of the FCM algoritm

Figure 4.6 shows the clustered image with histogram. We can see the lesions cleary from the segmented images.

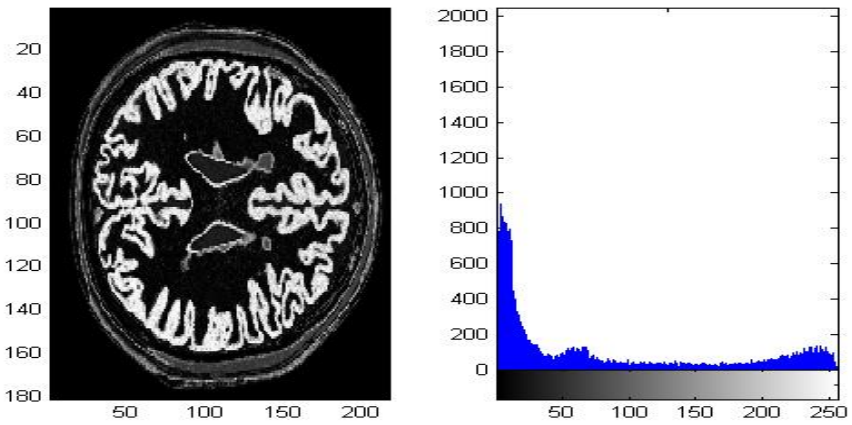


Figure 4.6 FCM clustered image and histogram.

However, if we apply a threshold to the image some of the CSF and gray matter parts are also seen as lesions. Figure 4.7 shows the thresholded result image. Figure 4.8 (a) shows original T1 image. Figure 4.8 (b) shows the original “gold” MS lesions image. Figure 4.8 (c) shows clustered image. Figure 4.8 (d) shows thresholded result image.

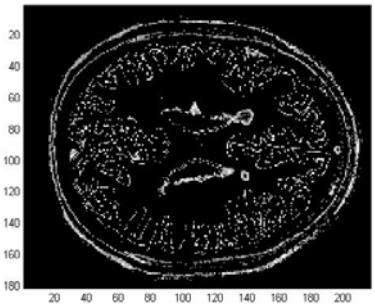


Figure 4.7 Thresholded image.

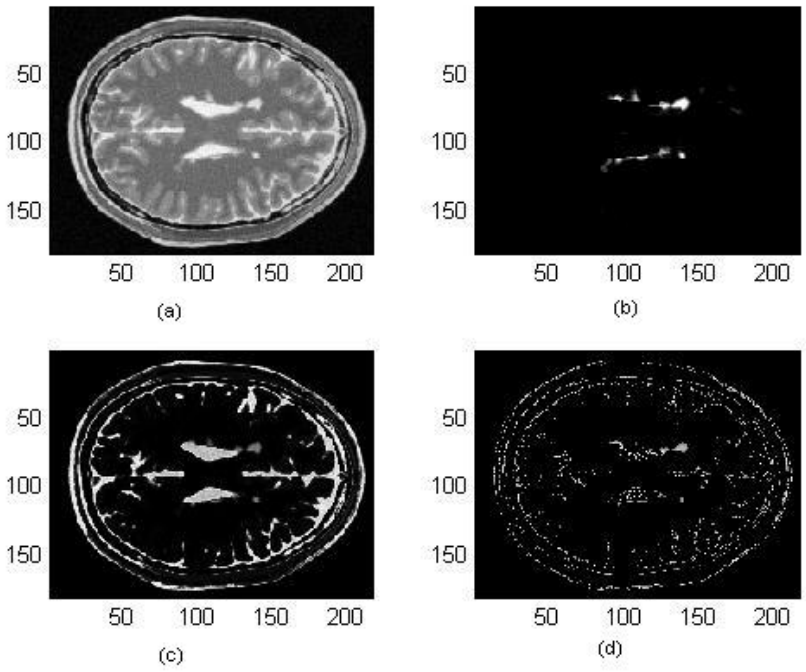


Figure 4.8 (a) Source T1 image slice number: 100 (b) MS lesions (c) Clustering with FCM image (d) Result image.

If we choose another slice (105) for same application, figure 4.9 shows the results.

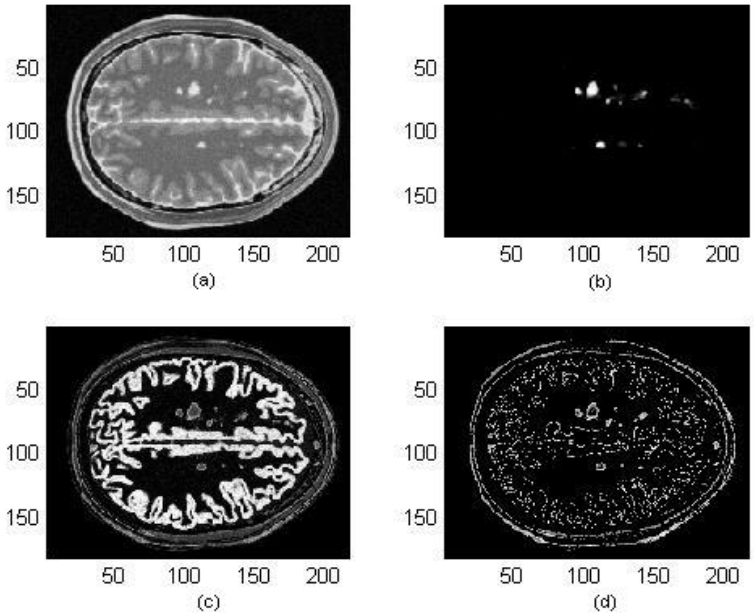


Figure 4.9 (a) Source T1 image slice number: 105 (b) MS lesions
(c) Clustering with FCM image (d) Result image.

4.4.2 K Means Clustering

K means clustering is a hard partition method. This means that, data are grupped in an exclusive way, so that if a certain data belongs to a definite cluster then it couldn't be included in another cluster. The algorithm for k-means clustering starts with computing the distances for all data points.

$$D_{ik}^2 = (x_k - v_i)^T (x_k - v_i), 1 \le i \le c, 1 \le k \le N \qquad \text{(Eq. 4.4)}$$

Where c is the number of clusters and first v_i is the random cluster center. Then calculation of the cluster centers is continous until $\prod_{k=1}^{n} \max \left| v^{(l)} - v^{(l-1)} \right| \neq 0$.

$$v_i^{(l)} = \frac{\sum_{j=1}^{N_i} x_i}{N_i} \qquad \text{(Eq. 4.5)}$$

Then partition matrix is calculated.

The algorithm for the k-means clustering is shown in the figure 4.10.

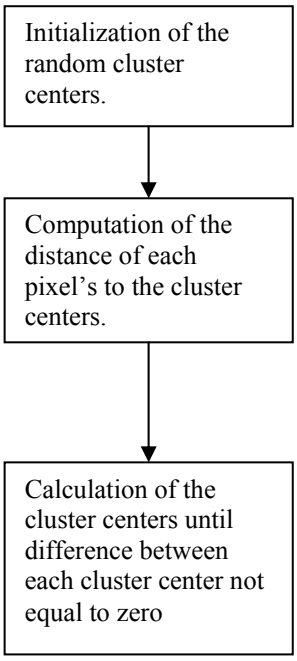


Figure 4.10 Algorithm for k-means clustering

For implementation of the MRI image to this algorithm, same slices (100, 105) are chosen.

K-means clustering is accomplished using MATLAB "fuzzy clustering toolbox" (Balasko B., 2006). In the application stage, first of all, 100. slice has been used for clustering. Input data for the clustering is non-linear combination of the T1-T2-PD images (Özkurt A., 2007). Some of the results are shown in the figure 4.11.

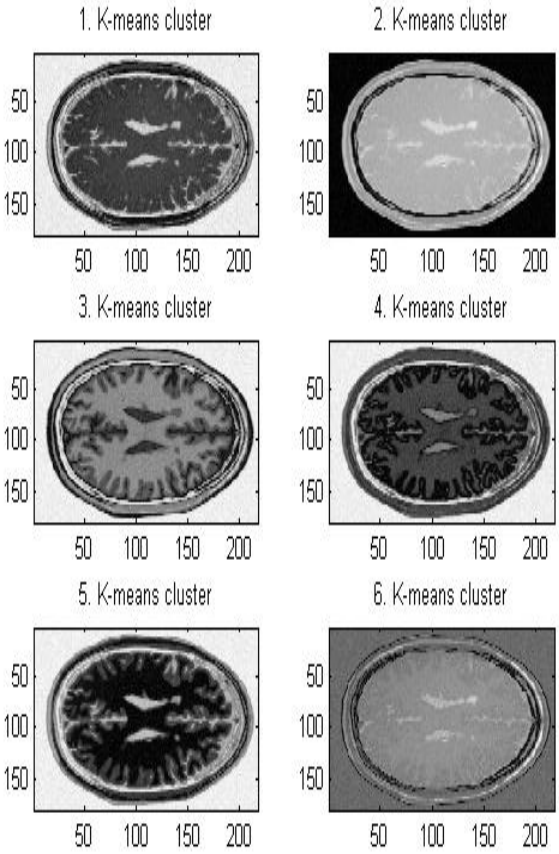


Figure 4.11 Some of the clustered images with k-means algorithm.

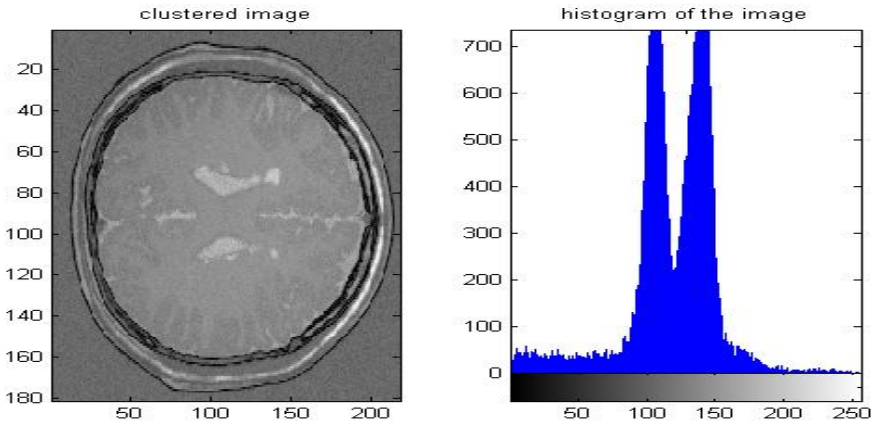


Figure 4.12 A clustered image with it's histogram.

Figure 4.12 shows the clustered image and histogram.

After thresholding for seperate specified areas, lesions and CSF areas didn't appear clearly. Figure 4.13 (d) shows the result thresholded image.

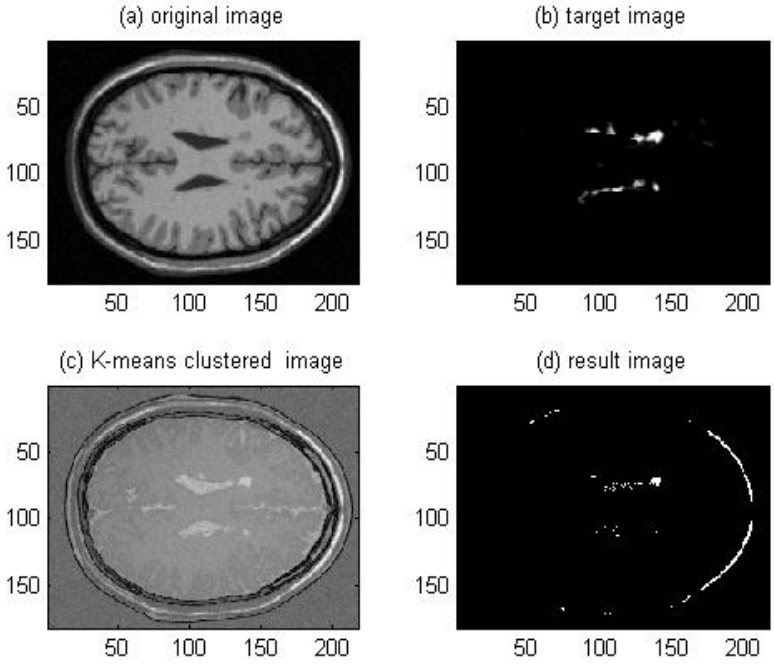
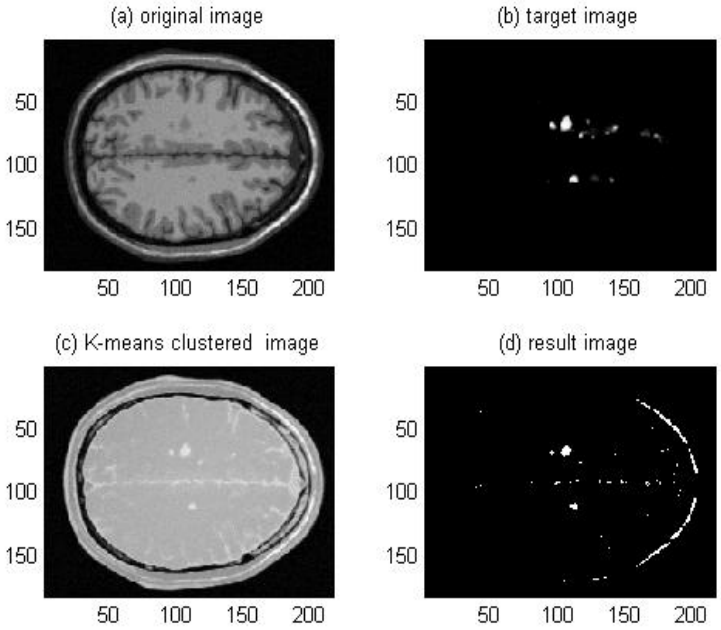


Figure 4.13 (a) Source T1 image slice number: 100 (b) MS lesions
(c) Clustering with K-means image (d) Result image.

Application for slice 105, has given the similar result as slice 100. CSF, lesions and skull didn't seperate clearly. Figure 4.14 (a) shows original T1 image. Figure 4.14 (b) shows the original "gold" MS lesions image. Figure 4.14 (c) shows clustered image. Figure 4.14 (d) shows thresholded result image.



.

Figure 4.14 (a) Source T1 image slice number: 105
(b) MS lesions
(c) Clustering with K-means image
(d) Result image.

4.4.3 K- Medoid Clustering

K-Medoid clustering is different from K-Means clustering. In K-Medoid clustering, the cluster centers are the nearest objects to the mean of data in one cluster.

$$V = \langle v_i \in X \,|\, 1 \le i \le c \rangle \qquad \text{(Eq. 4.6)}$$

Algorithm starts with the same as Eq. 4.4. Then original K-means cluster centers are calculated.

$$v_i^{(l)*} = \frac{\sum_{j=1}^{N_i} x_i}{N_i} \qquad \text{(Eq. 4.7)}$$

Nearest data points to be the cluster centers are chosen.

$$D_{ik}^{2*} = (x_k - v^*_i)^T (x_k - v^*_i) \qquad \text{(Eq. 4.8)}$$

Calculation of the cluster centers is continous until $\prod_{k=1}^{n} \max \left| v^{(l)} - v^{(l-1)} \right| \neq 0.$

Then partition matrix is calculated. Algorithm for the k-medoid clustering is shown in figure 4.15.

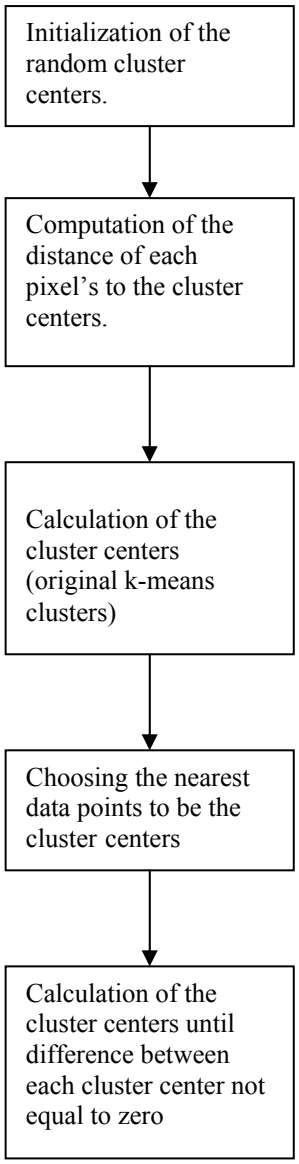


Figure 4.15 Algortihm for k-medoid clustering

For implementation of the MRI image to this algorithm, same slices (100 and 105) are choosen.

K-means clustering is accomplished using MATLAB "fuzzy clustering toolbox" (Balasko B., 2006). In the application stage, first of all, 100. slice has been used for clustering. Input data for the clustering is non-linear combination of the T1-T2-PD images (Özkurt A., 2007).

Some of the results are shown in figure 4.16.

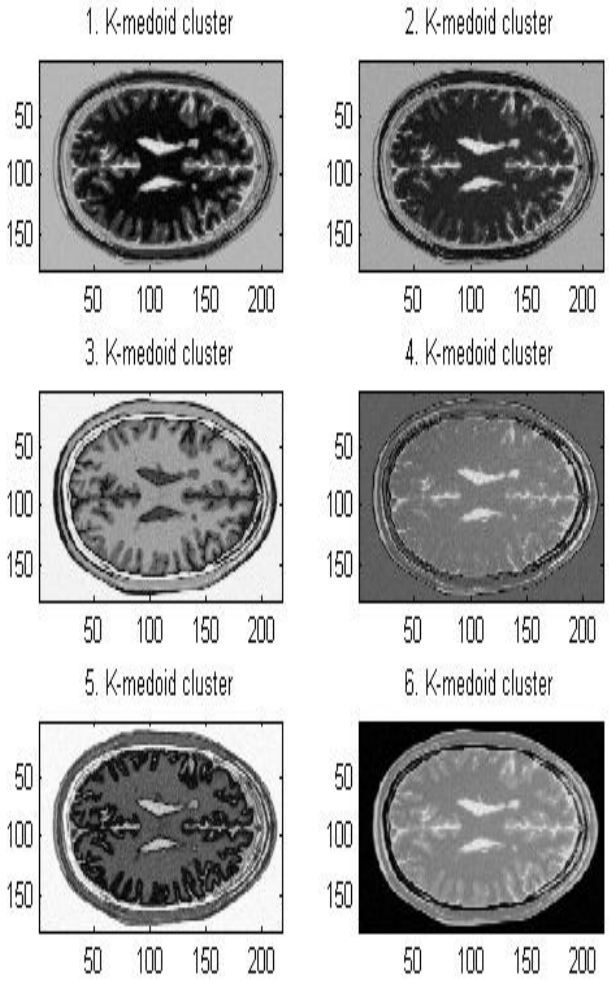


Figure 4.16 Some of the clustered images with k-medoid algorithm.

Figure 4.17 shows one of the clustered image and it's histogram.

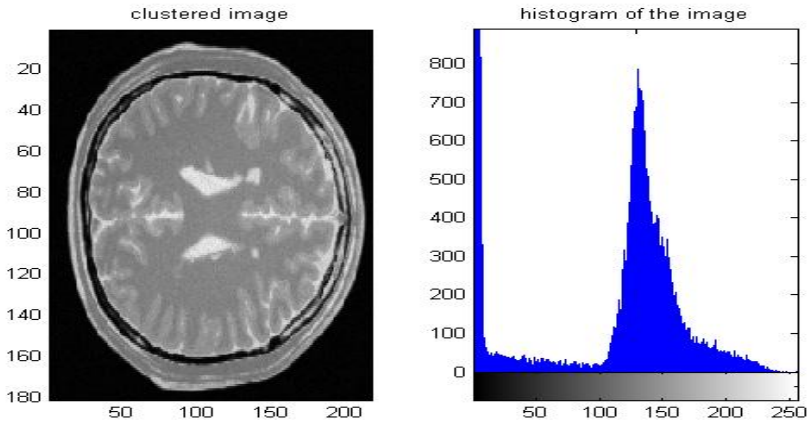


Figure 4.17 Image that clustered with K-medoid algorithm and it's histogram.

After thresholding for seperate specified areas, lesions and CSF areas didn't appear clearly. . Figure 4.18 (a) shows original T1 image. Figure 4.18 (b) shows the original "gold" MS lesions image. Figure 4.18 (c) shows clustered image. Figure 4.18 (d) shows thresholded result image

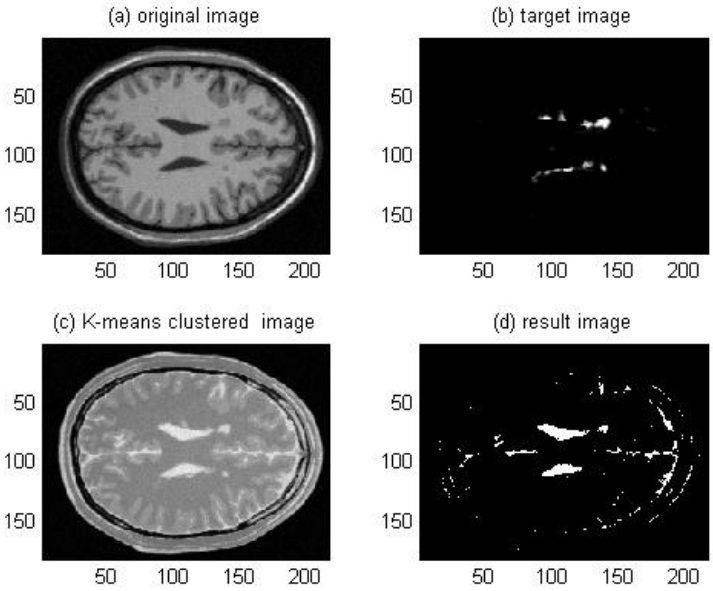


Figure 4.18 (a) Source T1 image slice number: 100
(b) MS lesions
(c) Clustering with K-medoid image
(d) Result image.

4.5 Principle Component Analysis (PCA)

Principle Component Analysis is the projection of the n-dimensional data into a lower q-dimensional data. First, the d-dimensional mean vector μ and dxd covariance matrix Σ are computed for all data sets. Next, the Eigenvectors and Eigenvalues are computed, and they are sorted according to decreasing Eigenvalue. The Eigenvector associated with the largest Eigenvalue has the same direction as the first principle component.

Covariance of the data set is:

$$F = \frac{1}{N}(x_k - v)(x_k - v)^T \qquad \text{(Eq. 4.9)}$$

N is the number of objects in the data set.

v is the mean of the data.

Principle component analysis is based on the projection of correlated high dimensional data onto a hyperplane (Tipping M. E., 1999). For mapping, to reduce the dimensionality of the data set, the first nonzero Eigenvalues and the corresponding Eigenvectors of the data set are used. F is the covariance matrix that includes the Eigenvalues in its diagonal in decreasing order. The U_i matrix includes the eigenvectors corresponding to the eigenvalues in its columns. The vector $y_{i,k}$ is a q dimensional reduced representation of the observed vector x_k.

$$y_{i,k} = W_i^{-1}(x_k) = W_i^T(x_k) \qquad \text{(Eq. 4.10)}$$

W_i is the weight matrix that contains the q principle orthonormal axes in its column.

$$W_i = U_{i,q}\Lambda_{i,q}^{\frac{1}{2}} \qquad \text{(Eq. 4.11)}$$

The algorithm for principle component analysis is shown in figure 4.19. This projection method is also known as Kahunen-Loeve Transform.

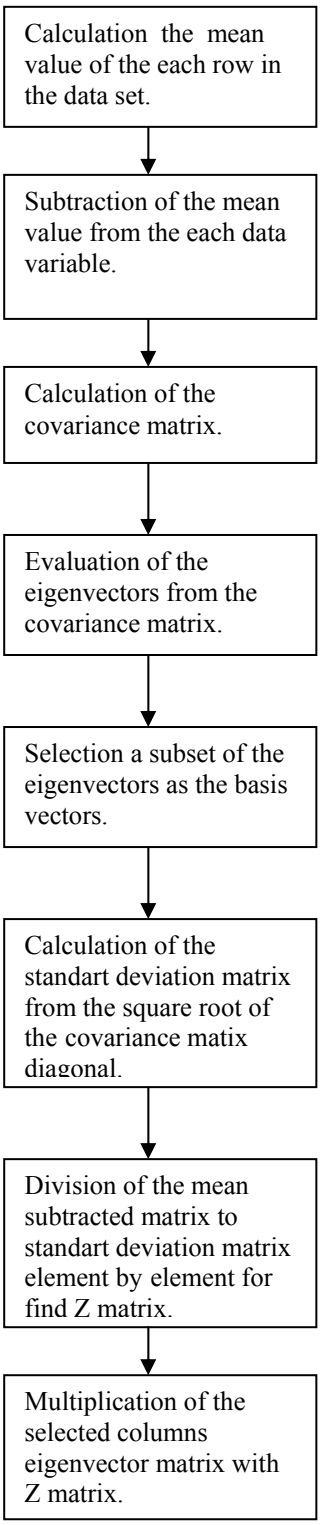


Figure 4.19 Algorithm of PCA

As we can see from the section 4.4.1, 4.4.2 and 4.4.3 clustering methods are not enough segment the lesions. The combination of the Principle Component Analysis with the overlapping (FCM) and exclusive (K-means and K-medoid) clustering methods, gives more succesfull results. PCA clustering is accomplised by MATLAB with "fuzzy clustering toolbox" (Balasko B., 2006).

Result segmented images are compared with the real MS lesions images and evaluated with two different methods.

The first comparing method is comparing the real MS lesions images and result segmented images pixels and finding false positive and negative pixels. Ratio of the true located pixels of the result image to the lesion area pixels of the target image gives the percentage of the performance.

The MATLAB code for calculation of the percentage of the true pixels is below:

```
testim=zeros(m,n);
% b is the target image
% result is the segmented image
for i=1:m;
   for j=1:n;
      if b(i,j)==result2(i,j);
         testim(i,j)=0;
         else testim(i,j)=255;
      end

   end
end
% testim image gives difference between target
% and result images
%calculate error
ms=length(find(b==255))%target pixel number
```

```
tp=length(find((b==255)&(result2==255)))

w1=ms-tp %false negative pixels
ts=length(find(testim==255))
rs=length(find(result2==255))
w2=rs-tp %false positive pixels
 %percent of true pixels
tt=(tp/ms)*100
```

The second method is root mean square error (RMSE) (Liu Zheng , 2006).

$$RMSE = \sqrt{\frac{\sum_{m=1}^{M}\sum_{n=1}^{N}\left[R(m,n)-I(m,n)\right]^2}{M \times N}} \qquad \text{(Eq. 4.12)}$$

Where, $R(m,n)$ is the referance image and $I(m,n)$ is the target image. $M \times N$ is the size of the images.

4.5.1 FCM With PCA

Application of the Fuzzy C Means clustered images together with the Principle Component Analysis and using the same data sets gives more specific segmentation results for lesions.

Figure 4.20 shows the results of the FCM plus PCA algorithm after 61 FCM iteration where the input of the PCA algorithm is data set from FCM input. Figure 4.21 shows the selected clustered image and it' s histogram. Figure 4.22 (a) shows the original T1 image. Figure 4.22 (b) shows the original MS lesions. Figure 4.22 (c) shows the segmented image by using fuzzy c-means and principle component analysis algorithms. Figure 4.22 (d) shows the thresholded result lesions of slice 100 of data set.

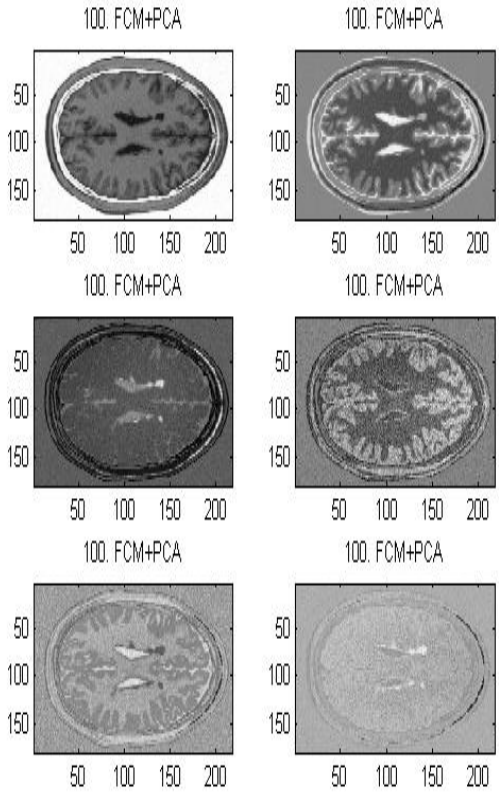


Figure 4.20 Different of the clustered results of the Fuzzy C-Means clustering with Principle Component Analysis for slice 100.

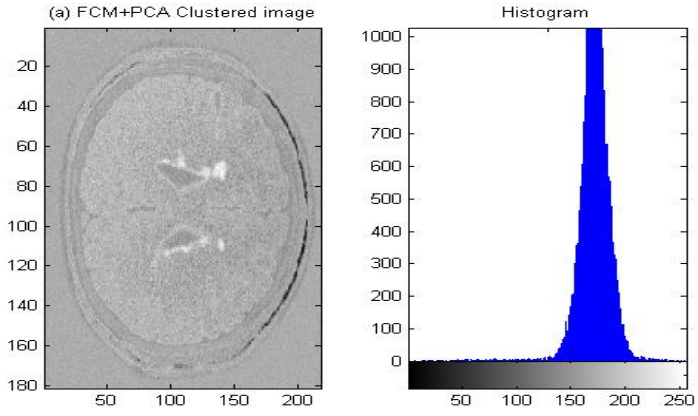


Figure 4.21 (a) FCA+PCA clustered image
(b) Histogram of the clustered image

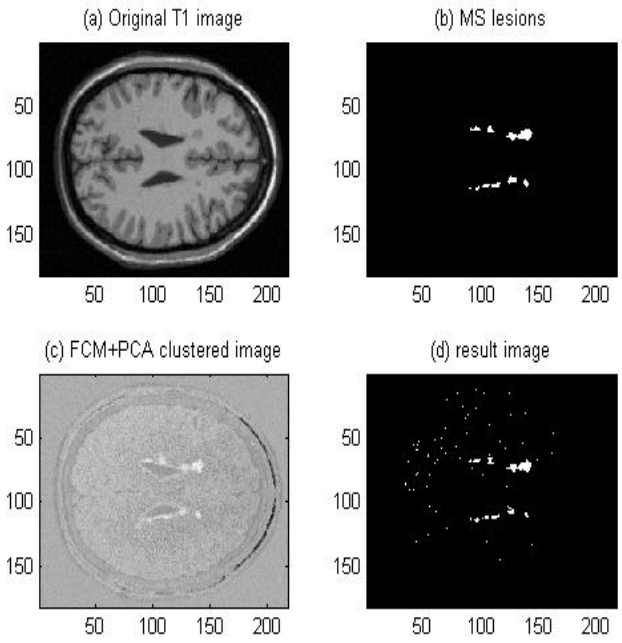


Figure 4.22 (a) Original image for slice 100
(b) MS lesions
(c) Segmented image with FCM and PCA
(d) Result Image

Figure 4.23 shows the threshold value according to true vs false pixels.

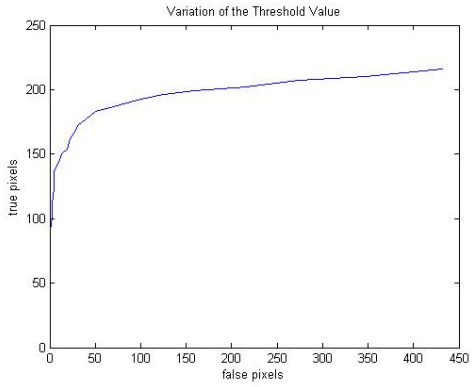


Figure 4.23 The variation of threshold value true pixels vs false pixels.

After 61 FCM iteration and PCA operation, comparision of the MS lesions image and result image , we can find 74.70% of lesion areas of slice 100. RMSE=0.86

According to histogram and threshold curve, the threshold value is selected 205 which gives the optimal true pixel and false positive pixel.

Figure 4.24 shows the same application's results for slice 105.

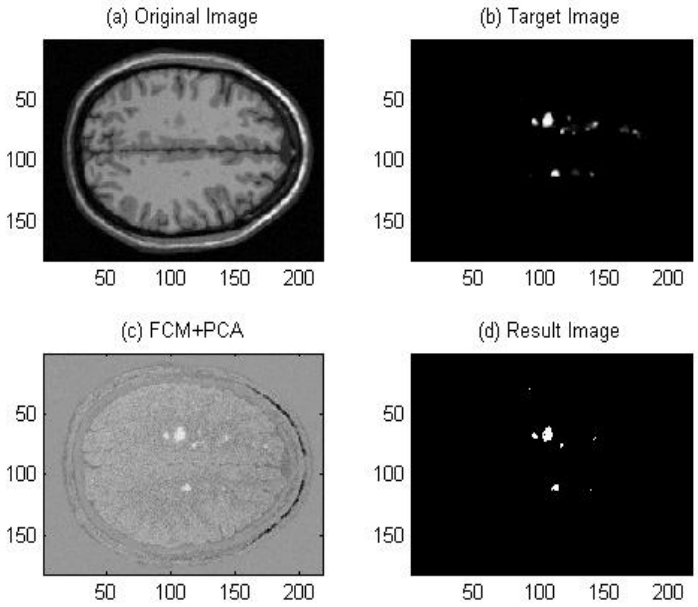


Figure 4.24 (a) Original image for slice 105
(b) MS lesions
(c) Segmented image for FCM and PCA algorithm
(d) Thresholded Image

Figure 4.25 shows the graphic of the variation of threshold value vs true pixels and false pixels.

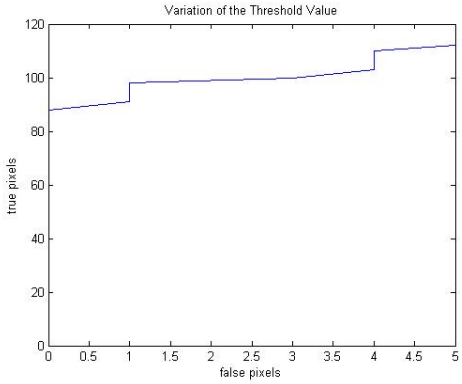


Figure 4.25 The variation of threshold value true pixels vs false pixels.

After 268 FCM iteration and PCA operation, comparision of the MS lesions image and result image , we can find 52.52% of lesion areas of slice 105. RMSE=0.72

Figure 4.26 shows the same applications results for slice 110. After 79 FCM iteration and PCA operation, comparision of the MS lesions image and result image , we can find 81.08% of lesion areas of slice 110. RMSE= 0.9.

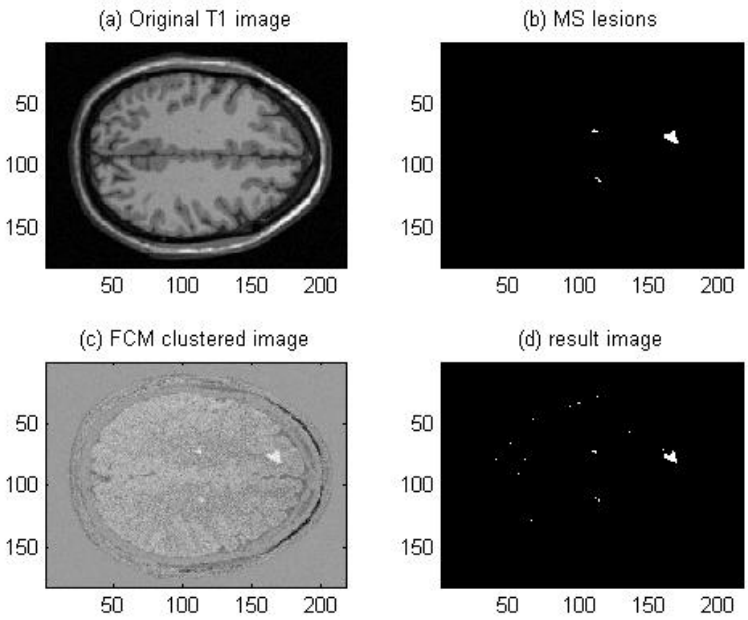


Figure 4.26 (a) Original image for slice 110 (b) MS lesions
(c) Segmented image for FCM and PCA algorithm (d) Result Image

Figure 4.27 shows the graphic of the variation of threshold value vs true pixels and false pixels.

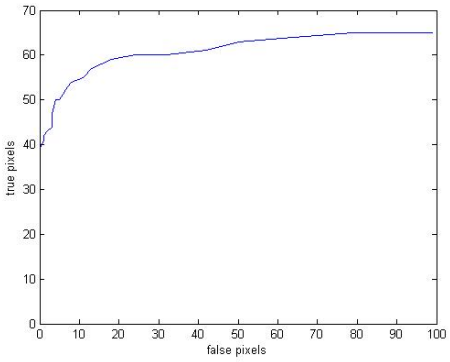


Figure 4.27 The variation of threshold
curve true pixels vs false pixels.

4.5.2 K-Means With PCA

The combination of K-means and PCA also gives more accurate results. Application of the K-Means clustered images together with the Principle Component Analysis and using the same data sets gives more specific segmentation results for lesions.

Figure 4.28 shows clustered images. Figure 4.29 shows segmented image and the histogram for finding threshold value.

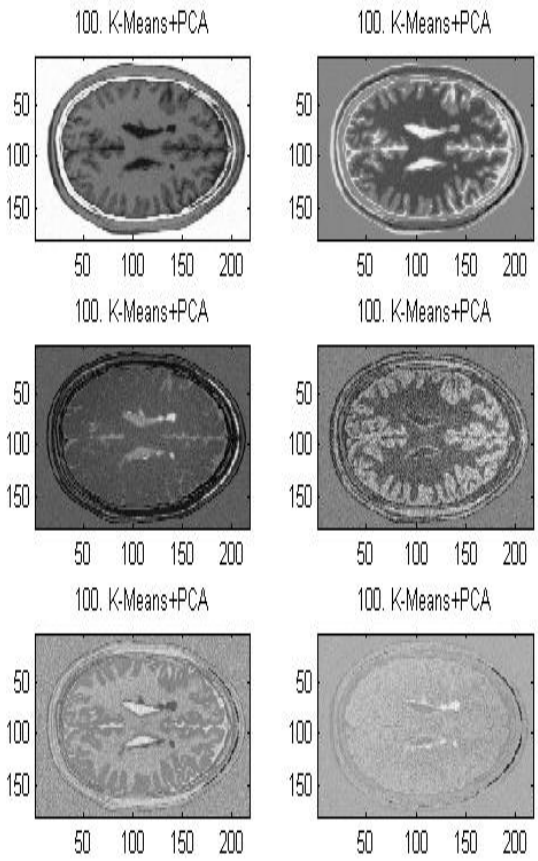


Figure 4.28 Some of the clustered results of the K-Means clustering with Principle Component Analysis for slice 100.

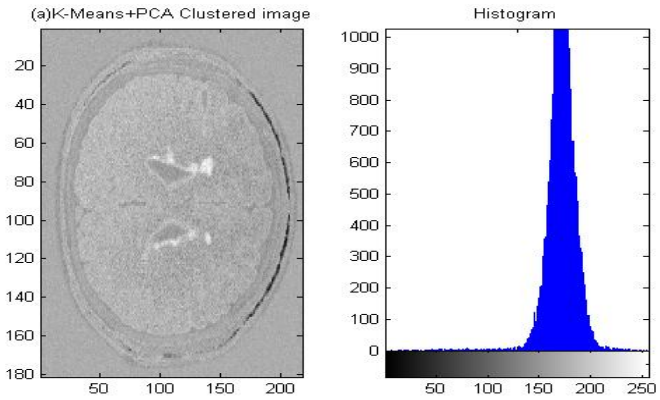


Figure 4.29 (a) K-Means+PCA segmented image
(b) Histogram

Figure 4.30 shows the graphic of the variation of threshold value vs true pixels and false pixels. According to histogram and threshold graphic, the threshold value is selected 205.

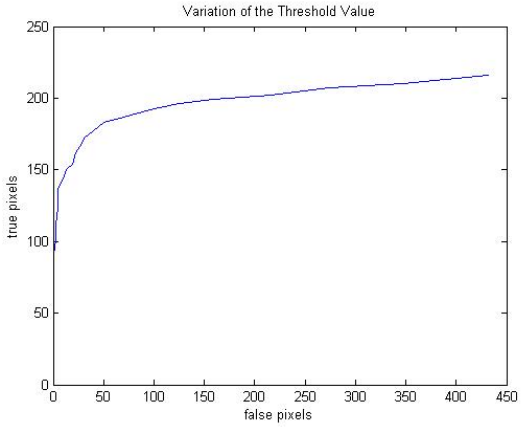


Figure 4.30 The variation of threshold value vs true pixels and false pixels.

Figure 4.31 shows result images.

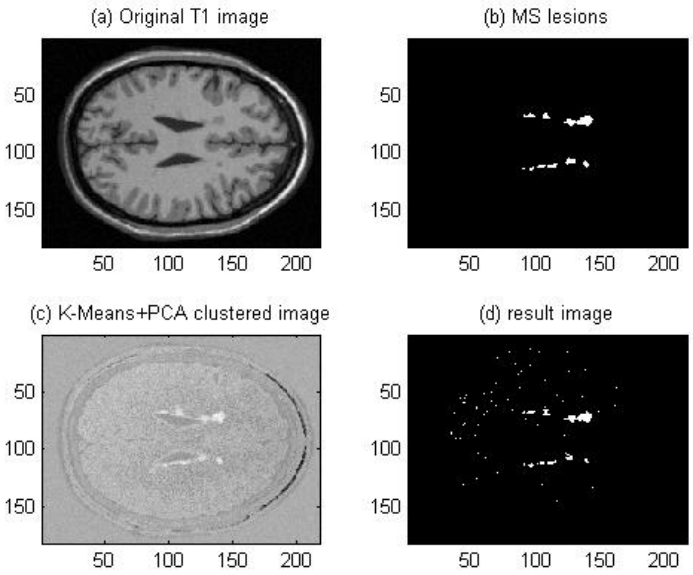


Figure 4.31 (a) original T1 image (b) MS lesions.
(c) Clustered image with K-means and PCA.
(d) Result image.

Comparision of the MS lesions image and result segmented image , we can find 74.70% of lesion areas. RMSE is 0.86.

Figure 4.32 shows the results for same application to slice 105 , we can find 52.2% of lesion areas. RMSE is 0.72.

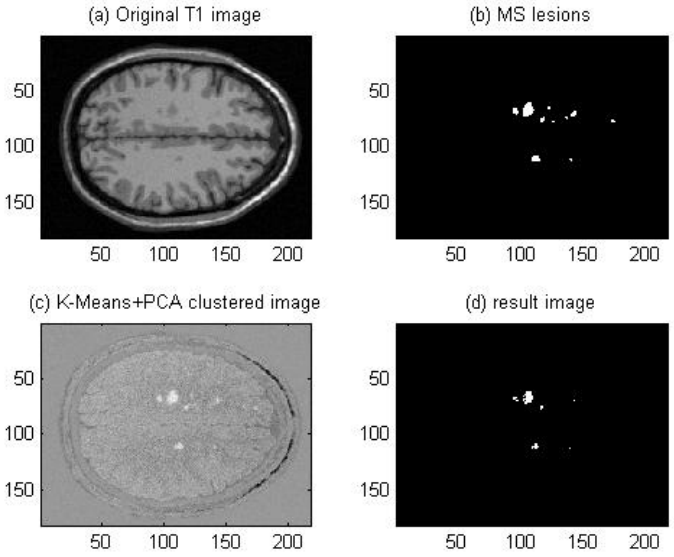


Figure 4.32 (a) original T1 image for slice 105. (b) MS lesions .
(c) Clustered image with K-means and PCA. (d) Result image.

Figure 4.33 shows the histogram of the selected cluster of slice 105

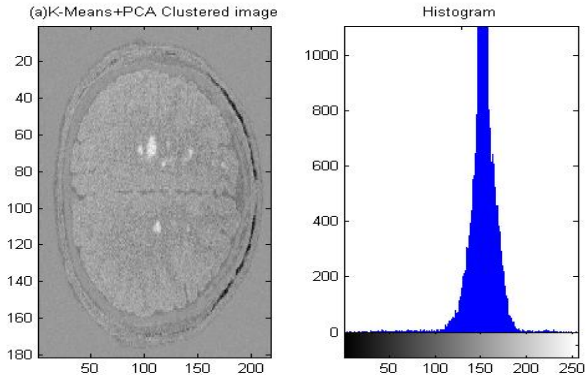


Figure 4.33 (a) K-Means+PCA segmented image
(b) Histogram

Figure 4.34 shows the graphic of the variation of threshold value vs true pixels and false pixels. According to histogram and threshold graphic, the threshold value is selected 205.

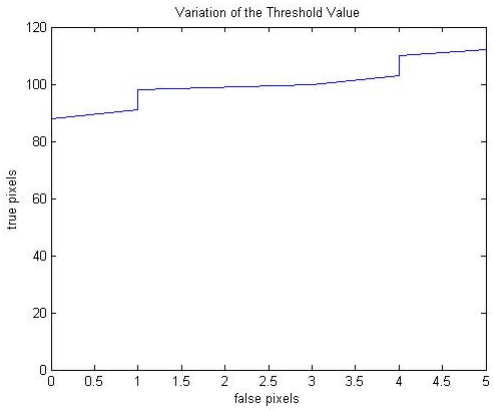


Figure 4.34 The variation of threshold value true pixels vs false pixels.

Figure 4.35 shows the results for same application to slice 110 , we can find 81.08% of lesion areas. RMSE is 0.9.

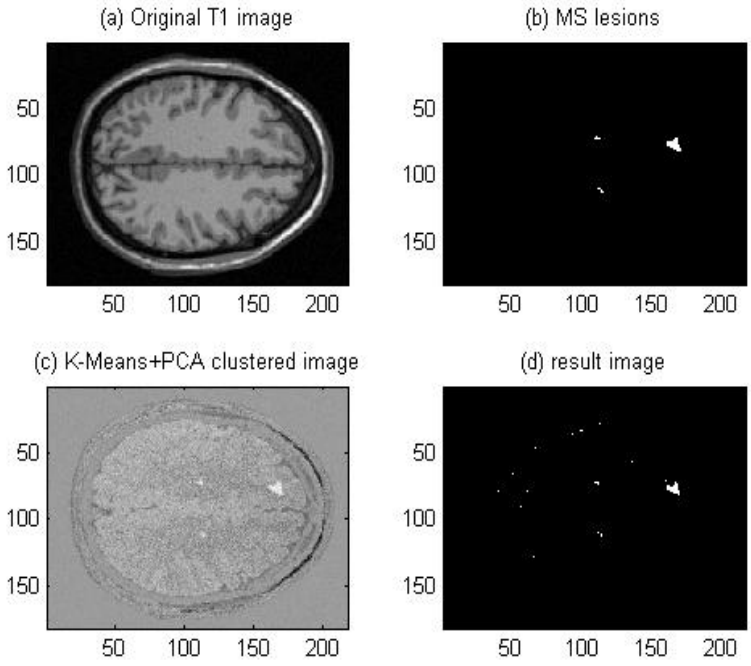


Figure 4.35 (a) Original T1 image for slice 110
(b) Original MS lesions.
(c) Clustered image with K-means and PCA.
(d) Result thresholded image.

Figure 4.36 shows the histogram of the selected cluster of slice 110.

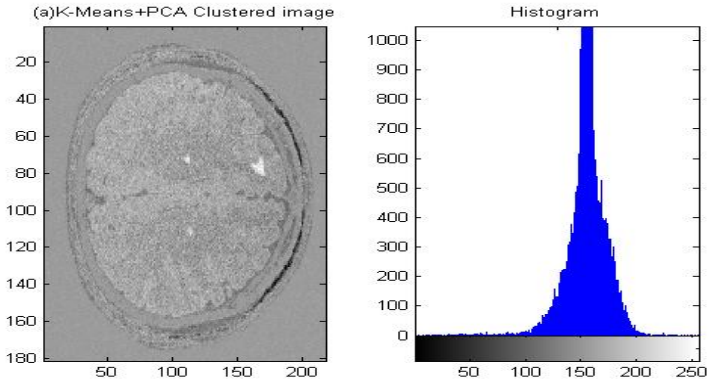


Figure 4.36 (a) K-Means+PCA segmented image
(b) Histogram

Figure 4.37 shows the graphic of the variation of threshold value, true pixels vs false pixels. According to histogram and threshold graphic, the threshold value is selected 205.

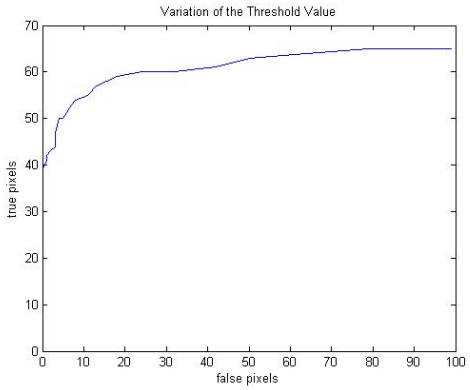


Figure 4.37 The variation of threshold value true pixels vs false pixels.

4.5.3 K-Medoid With PCA

The combination of K-medoid and PCA and the results are below. Figure 4.38 shows clustered images. K-medoid results are very similar to the K-means results lesion clusters as can be clearly seen. The difference between k-means and k-medoid is in the calculation of the cluster centers.

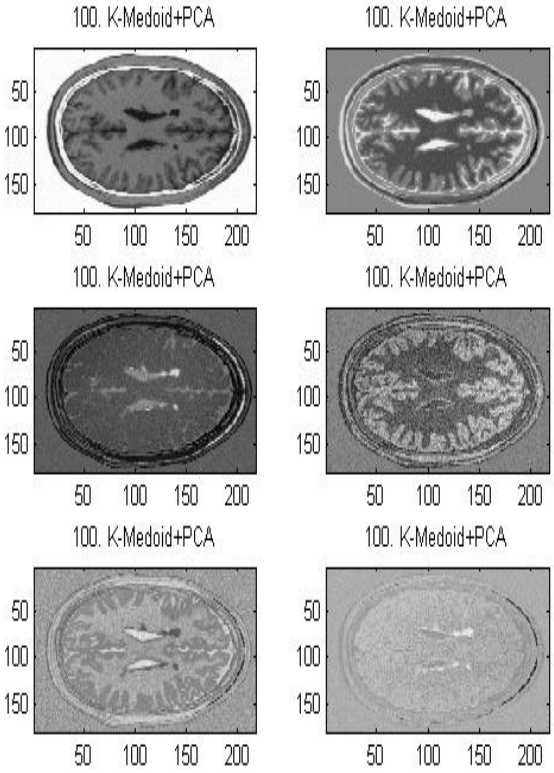


Figure 4.38 Some of the results for combination of the K-medoid and PCA Segmentation.

Figure 4.39 shows the selected clustered image and histogram.

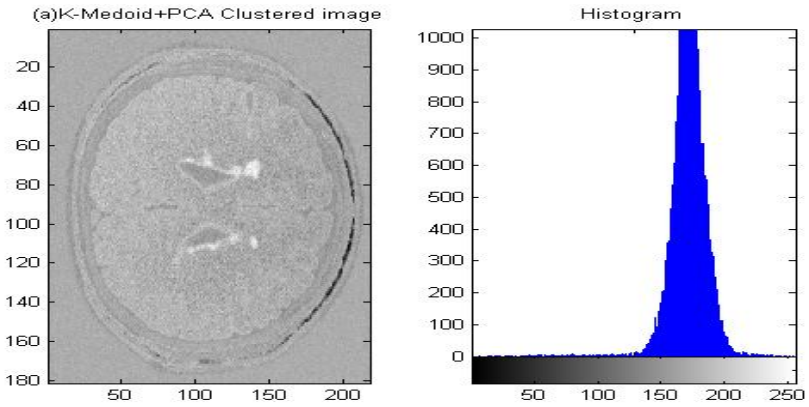


Figure 4.39 (a) K-Medoid+PCA segmented image (b) Histogram

Figure 4.40 shows the graphic of the variation of threshold value, true pixels vs false pixels. According to histogram and threshold graphic, the threshold value is selected 205.

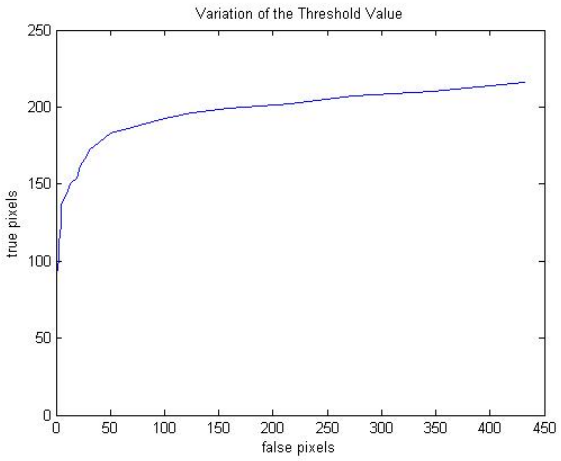


Figure 4.40 The variation of threshold value , true pixels vs false pixels.

Figure 4.41 (a) shows original T1 image. Figure 4.41 (b) shows the original “gold” MS lesions image. Figure 4.41 (c) shows clustered image. Figure 4.41 (d) shows thresholded result image

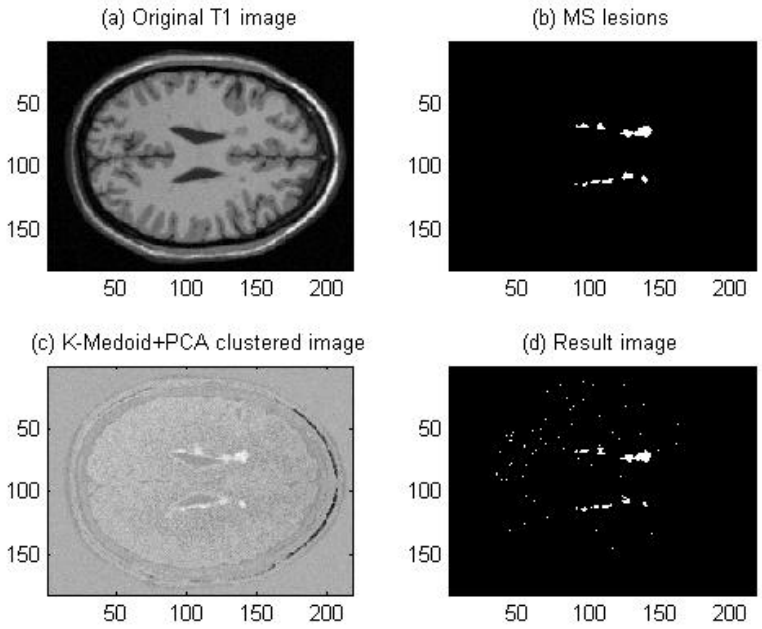


Figure 4.41 (a) original T1 image for slice 100. (b) MS lesions
(c) Clustered image with K-medoid and PCA. (d) Result image.

Comparision of the target image and result image , we can find 74.70% of lesion areas. RMSE is 0.86.

Figure 4.42 shows the results for same application to slice 105, we can find 52.2% of lesion areas. RMSE is 0.72.

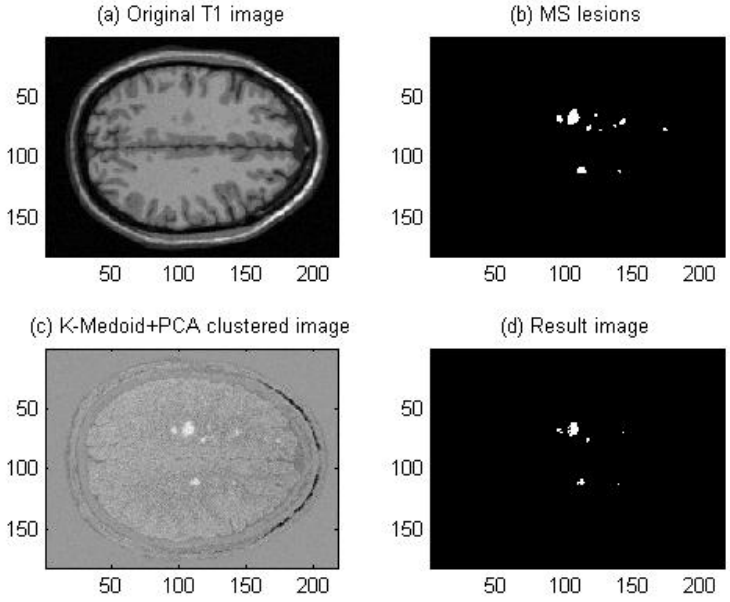


Figure 4.42 (a) original T1 image for slice 105. (b) MS lesions.
(c) Clustered image with K-medoid and PCA. (d) Result image.

Figure 4.43 shows the selected clustered image and histogram.

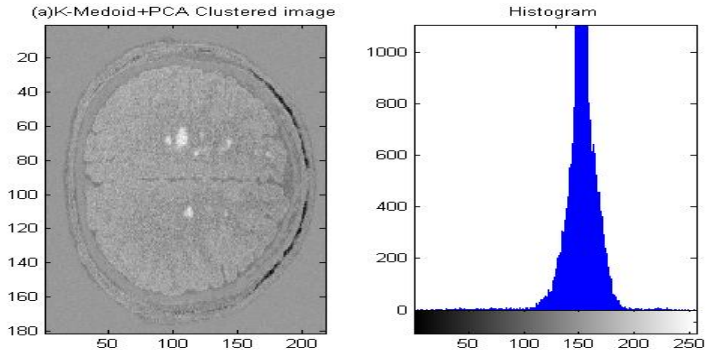


Figure 4.43 (a) K-Medoid+PCA segmented image (b) Histogram

Figure 4.44 shows the graphic of the variation of threshold value, true pixels vs false pixels. According to histogram and threshold graphic, the threshold value is selected 205.

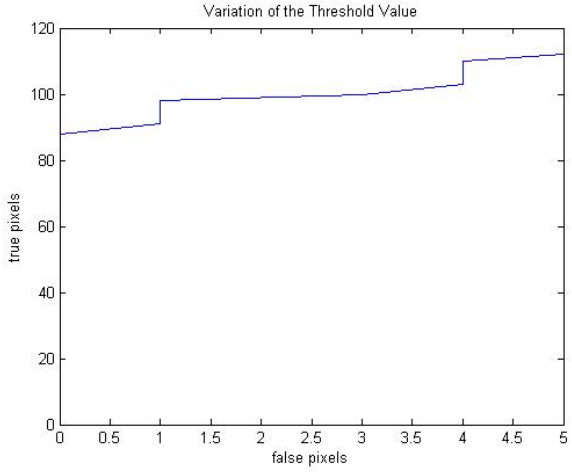


Figure 4.44 The variation of threshold value , true pixels vs false pixels.

Figure 4.45 shows the results for same application to slice 110, we can find 81.08% of lesion areas. RMSE is 0.9.

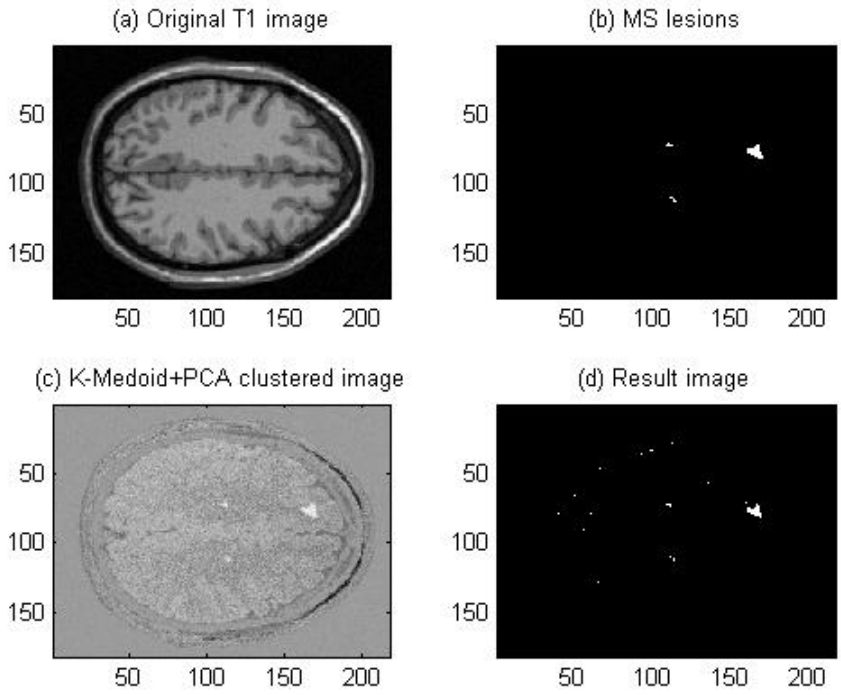


Figure 4.45 (a) original T1 image for slice 110. (b) MS lesions.
(c) Clustered image with K-medoid and PCA. (d) Result image.

Figure 4.46 shows the selected clustered image and histogram.

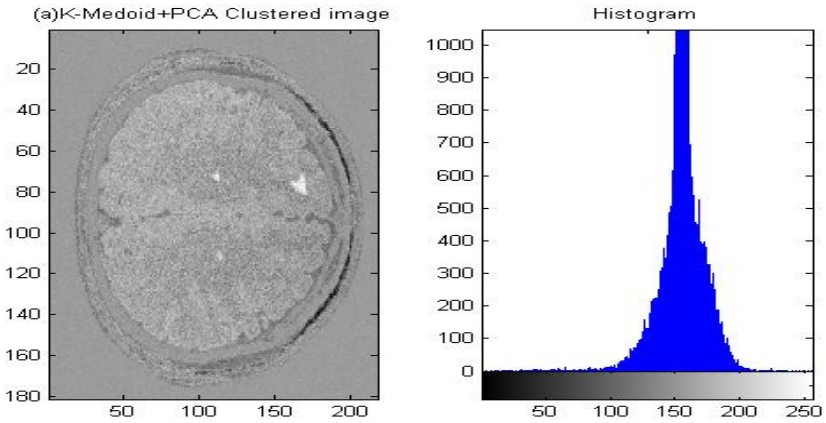


Figure 4.46 (a) K-Medoid+PCA segmented image (b) Histogram

Figure 4.47 shows the graphic of the variation of threshold value, true pixels vs false pixels. According to histogram and threshold graphic, the threshold value is selected 205.

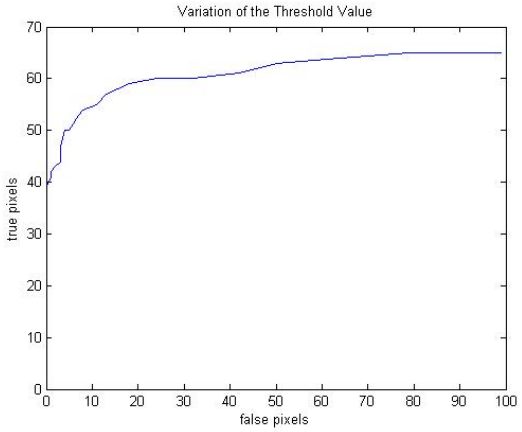


Figure 4.47 The variation of threshold value , true pixels vs false pixels.

4.6 Independent Component Analysis (ICA)

In the Independent Component Analysis (ICA), source signal s(t) is registered as x(t). Where $s(t) = [s_1(t), ..., s_n(t)]$ and $x(t) = [x_1(t), ..., x_m(t)]$. If a nonlinear interference appears, A is the interefence matrix between source and observed signals, $x(t) = As(t)$. The goal here is estimation of the source signals from the observed signals and find inverse interference matrix where $y(t) = Wx(t) = A^{-1}x(t)$. If there is no information about the sources, in a blind source seperation case, statistical independence of the independent component analysis sources and has no Gaussian probability density function asumptions can be considered (Özkurt A., 2007) .

The input bands are 24 different non-linear composition of the MRI images. In the application stage, first of all 100. slice has been used for clustering. Independent Component Analysis is accomplised by MATLAB ICALAB toolbox. Results are below.

We have 24 inputs and 24 different independent components. Figure 4.48 shows these results. Figure 4.49 shows the selected ICA clustered image and histogram.

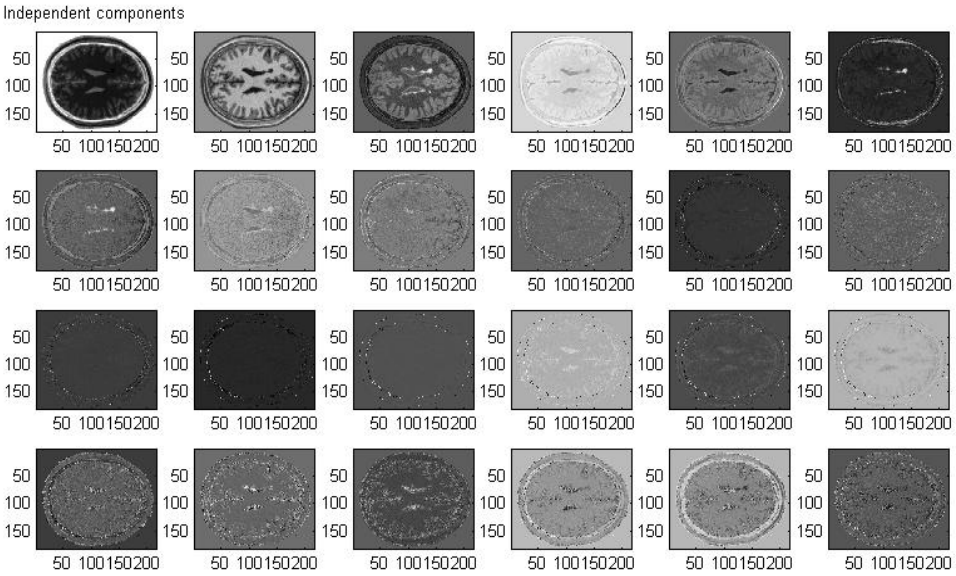
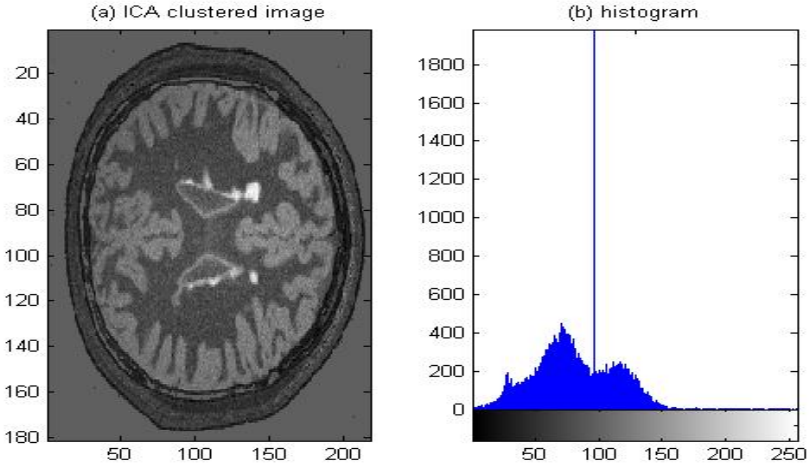


Figure 4.48 Independent components of the slice 100.



(a) (b)

Figure 4.49 (a) Independent Component Analysis Clustered Image
(b) Histogram

Figure 4.50 shows the graphic of the variation of threshold value, true pixels vs false pixels. According to histogram and threshold graphic, the threshold value is selected 154.

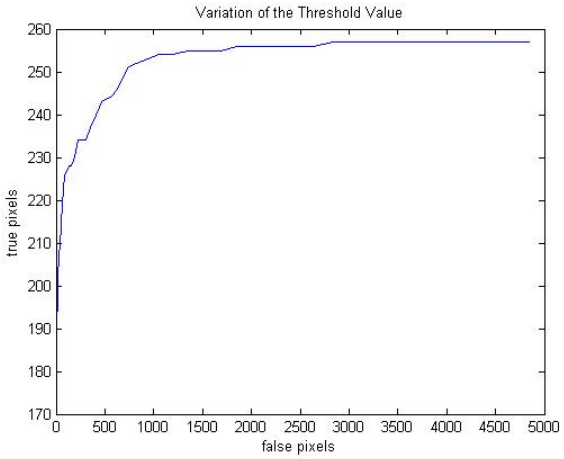


Figure 4.50 The variation of threshold value , true pixels vs false pixels.

Figure 4.51 shows the original T1 image and MS lesions images with ICA clustered and result thresholded image.

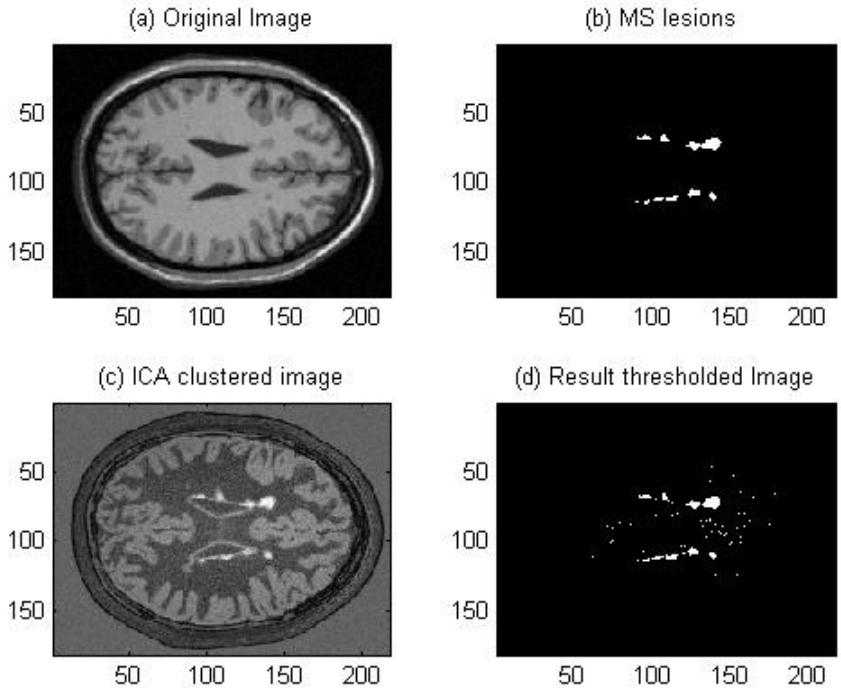


Figure 4.51 (a) Original T1 image (b) MS lesions.
(c) ICA clustered image (d) Result image

Comparision of the target image and result image , we can find 87.93 % of lesion areas. RMSE is 0.93.

Figure 4.52 shows the results for same application to slice 110 , we can find 54.05 %of lesion areas. RMSE is 0.73.

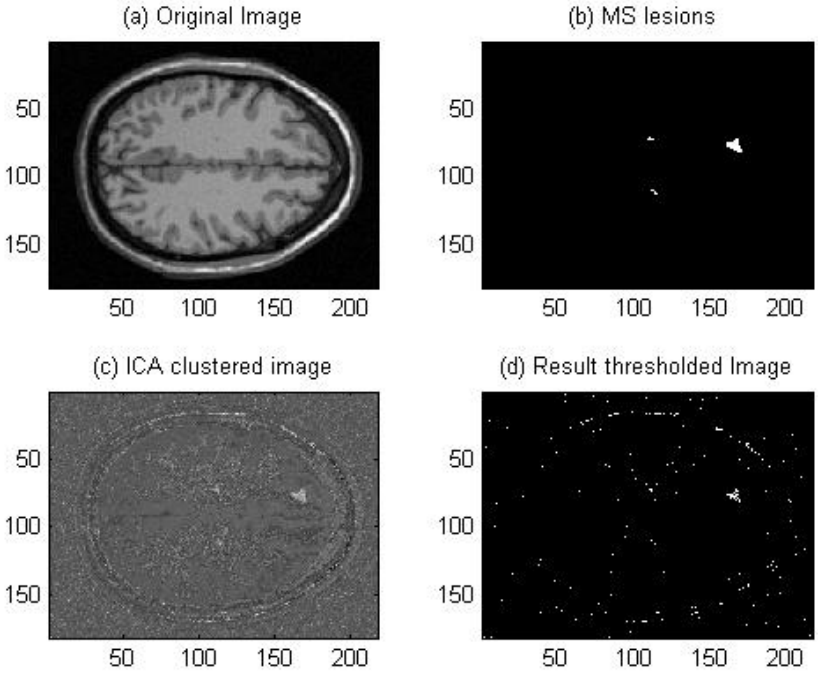


Figure 4.52 (a) Original T1 image (b) MS lesions
(c) ICA clustered image (d) Result image

Figure 4.53 shows the selected clustered image and histogram.

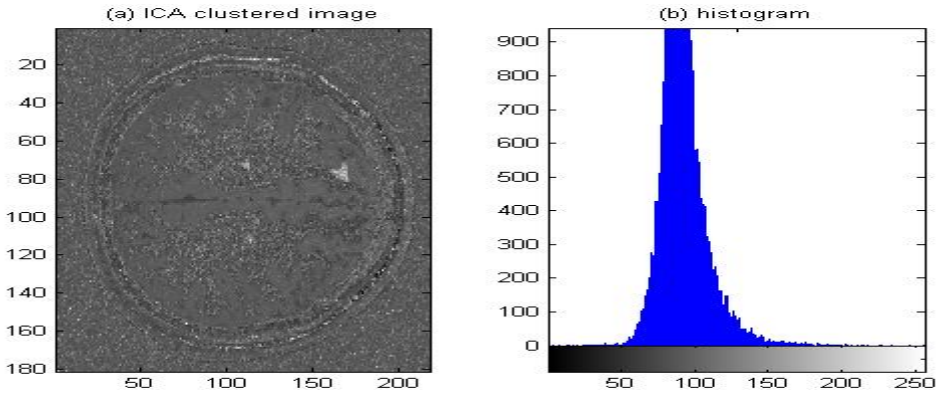


Figure 4.53 (a) Independent Component Analysis Clustered Image
(b) Histogram

Figure 4.54 shows the graphic of the variation of threshold value, true pixels vs false pixels. According to histogram and threshold graphic, the threshold value is selected 154.

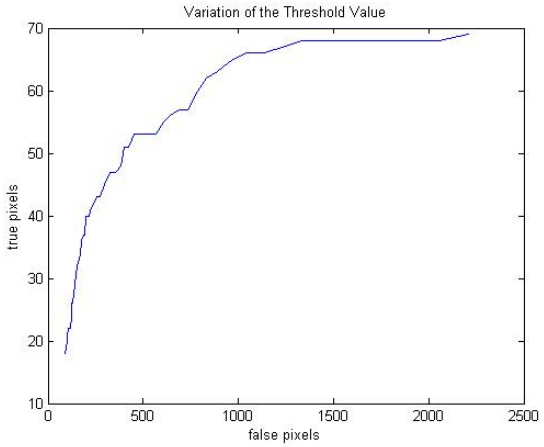


Figure 4.54 The variation of threshold value , true pixels vs false pixels.

Figure 4.55 shows the results for same application to slice 105 , we can find % 34.4444 of lesion areas. RMSE is 0.58.

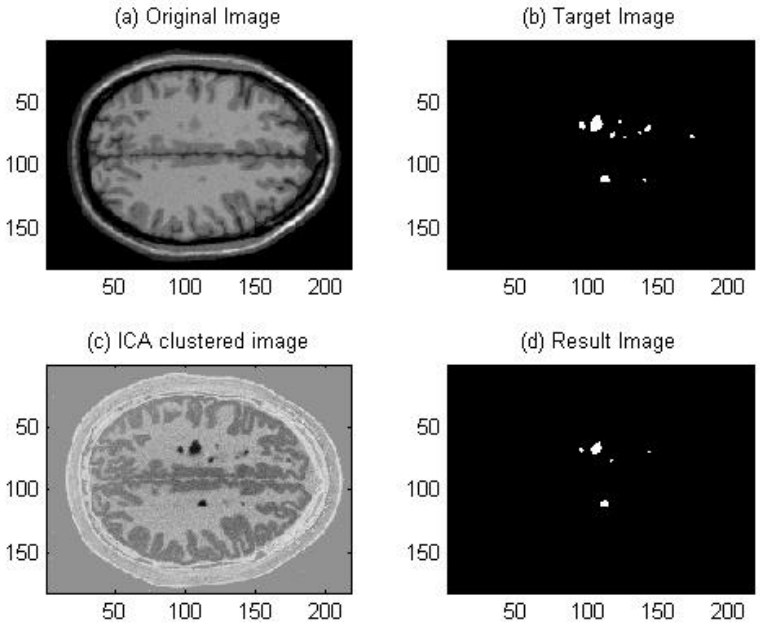


Figure 4.55 (a) Original T1 image (b) MS lesions (c) ICA clustered image (d) Result image

4.7 Segmentation With Neural Networks

There are many different approaches used to classify artificial neural network. For example, neural networks can be classified according to how they learn or the type of training that is required, the various application they perform. Those that use activation functions versus basis functions, whether they are recurrent or nonrecurrent, the type of training inputs. Neural networks can be classify three main class:

i) Single layer feedforward networks.

ii) Multilayer feedforward networks.

iii) Recurrent networks.

Multilayer feedforward networks have common and complex nerveous system. Typically have several levels of hidden layers. Partially connected are more common than fully connected. Multilayer feedforward networks are also have supervised learning rules.

4.7.1 Segmentation With Feedforward Multilayer Perceptron Algorithms

The first step for training the multilayer perceptron by using the standard backpropagation algorithm is initialize the network synaptic weights to small random values. The scaling factor for initialize the weights is depend to the number of component in input layers and number of neurons in hidden layer.

$$\gamma = 0.7\sqrt[n_0]{n_1} \qquad \text{(Eq. 4.13)}$$

n_0 is the number of components in the input layer

n_1 number of neurons in hidden layer.

$$w_{ij} = \gamma \frac{w_{ij}}{\sqrt{\sum_{i=1}^{n_1} w_{ij}^{\,2}}} \qquad \text{(Eq. 4.14)}$$

w_{ij} is weight for each neuron.

The second step is present an input pattern and calculate the network responce from the set of training input and output patterns.

The third step is comparision of the network response with the actual output of the network and compututions of local errors.

$$\delta_j^{(s)} = (d_{qh} - x_{out,j}^{(s)}) g(v_j^{(s)}) \quad \text{(Eq. 4.15)}$$

$\delta_j^{(s)}$ is the error for output layer.

d_q is the desired network output.

x_{out} is the actual output of the network.

v is the activity level for ith level.

$$\delta_j^{(s)} = \left(\sum_{h=1}^{n_{s+1}} \delta_h^{(s+1)} w_{hj}^{(s+1)} \right) g(v_j^{(s)}) \quad \text{(Eq 4.16)}$$

Equation 4.15 shows the error for hidden layers.

The fourth step is update of the network weights in according to equation 3.12.

$$w_{ji}^{(s)}(k+1) = w_{ji}^{(s)}(k) + \mu^{(s)} \delta_j^{(s)} x_{out,i}^{(s)} \quad \text{(Eq. 4.17)}$$

The last step is contiune the step two through four until the network reaches a predetermined level of accuracy in producing the adequate response for all the training patterns.

After training the system, the result weights can be used to simulate different input data.

Neural network segmentation with feedforward multilayer perceptron method is accomplised by MATLAB Neural Network Toolbox. Input data for the clustering is combination of T1-T2-PD images.

4.7.2 Application of Feedforward Multilayer PerceptronNeural Network

The input for training the network is the combination of T1, T2 and PD weighted images. The combination of these three images are feeded as input to the network. Training targets are chosen the target slice of the same MRI images. 16 input and 16 hidden layer and 1 output layer were used. As an activation function, tan-sigmoid transfer function was used. 400 epochs are chosen for the training epochs and after

system was trained, the network was simulated with the training data again. Figure 4.56 shows the performance curve and table 4.1 shows the training weights and bias values after 400 epoch.

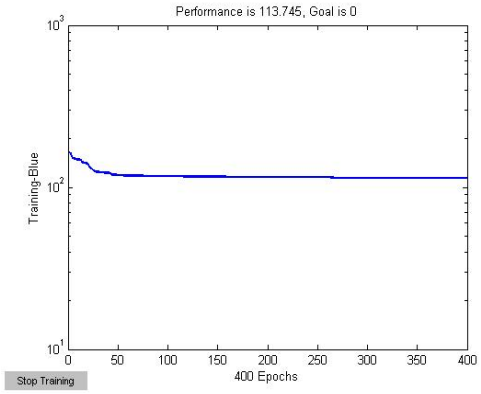


Figure 4.56 Training of the system with 400 epochs

Table 4.1 Neuron weights and bias values

Layer1_Weights			Layer1_Bias	Layer2_Weights	Layer2_Bias
0.6802	0.0658	0.6036	-7.7974	5.0769	0.2893
0.2828	-1.0518	-0.7050	7.3973	0.0046	
-0.2127	-0.1757	-0.5201	-1.1599	-0.5143	
-0.1727	-0.0396	0.1228	3.3188	14.8532	
-1.2199	-0.1999	-0.0644	-4.7986	-1.1221	
-1.2192	0.4724	0.7404	1.9998	11.0729	
0.3333	0.2548	-0.1285	6.6871	1.5080	
-2.5918	2.6181	1.2653	-4.7060	-0.1760	
0.1582	0.0636	-0.1590	-1.4747	-9.2721	
-1.4843	0.4197	0.5851	2.9028	5.0691	
-0.0865	-0.0826	0.0313	16.6073	-31.1109	
0.4582	-0.2269	-0.0372	-1.3992	1.3818	
-0.0722	-0.1426	-0.2334	-0.7747	0.0369	
0.1324	0.0969	-0.1407	-0.3431	18.4735	
-0.3896	0.6937	0.0892	1.0112	0.9157	
0.0410	0.2808	0.0887	1.0352	0.3738	

Figure 4.57 shows the training result and histogram.

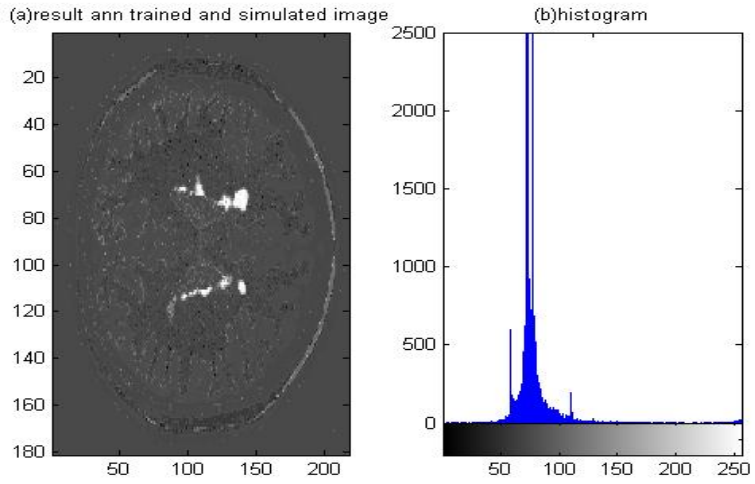


Figure 4.57 (a) Simulated image 100 (b) Histogram

According to the histogram from figure 4.56, lesion areas appear bright. After a simple thresholding result image is shown in figure 4.58 (a). Figure 4.58 (b) shows the target MS lesions image.

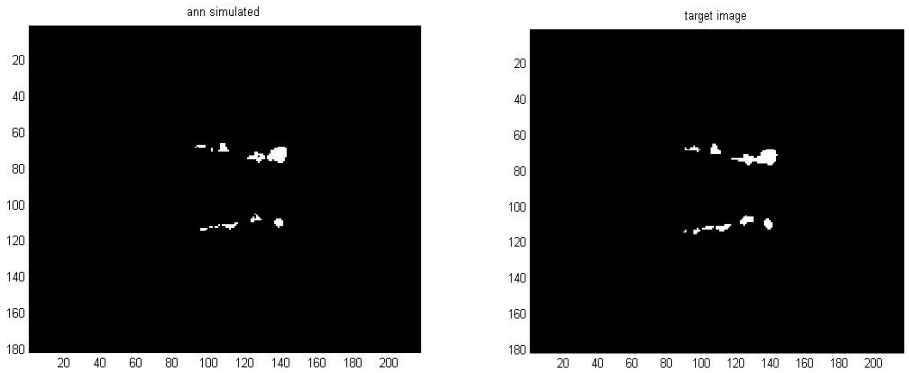


Figure 4.58 (a) Result of the simulated image of slice 100. (b) MS lesions

Calculation of the error between the ANN simulated image and target image:

With the comparision of the target MS lesions image and result image , we can find %81.3230 of the lesion areas. In according to Eq 4.7 root mean square error value is:

RMSE=0.9.

After training and simulating system, the same network was used to simulate different MRI image, slice 105 and 110. Figure 4.60 shows the simulated image 105 and it's histogram. Figure 4.59 shows the target MS lesions image of slice 105 and the thresholded result image. Figure 4.61 shows the simulated image 110 and it's histogram. Figure 4.62 shows the target image of slice 110 and the thresholded result image.

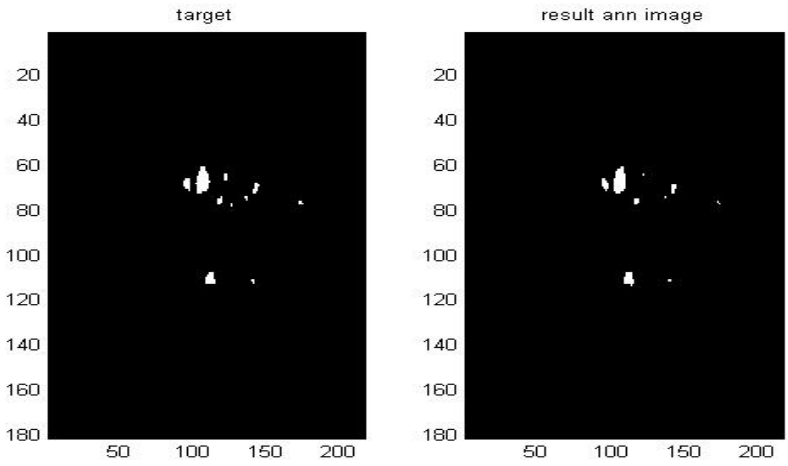


Figure 4.59 (a) MS lesions (b) result ann image of slice 105.

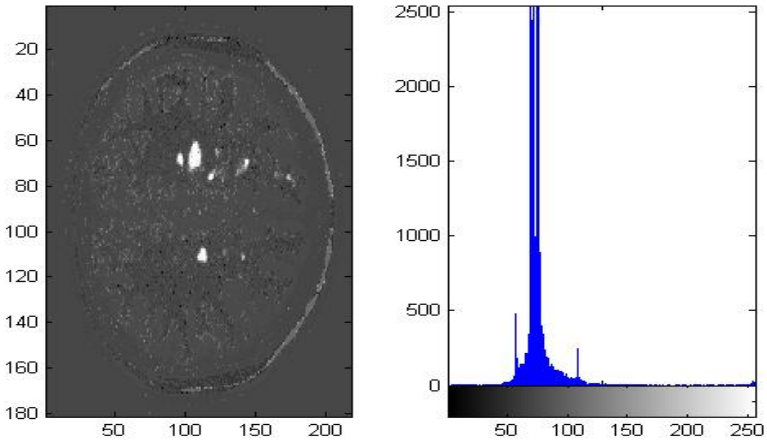
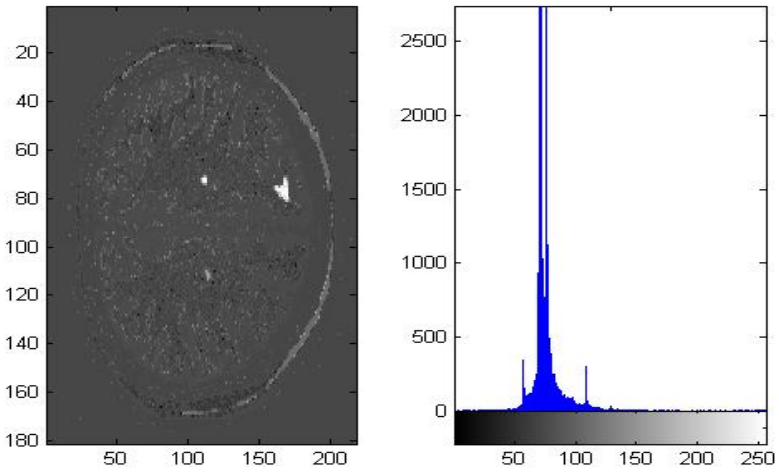


Figure 4.60 (a) ANN simulated image (slice 105) (b) Histogram

Comparision of the target image 105 and result image 105 , we can find %79.444 of the lesion areas. In according to Eq 4.7 root mean square error value is:

RMSE=0.89.



(a) (b)

Figure 4.61 (a) ANN simulated image (slice 110) (b) Histogram

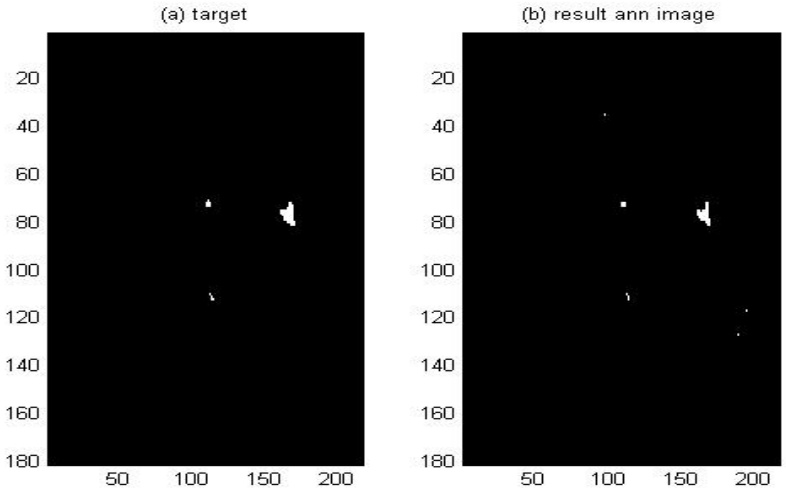


Figure 4.62 (a) MS lesions (b) result image of slice 110.

Comparision of the target image 110 and result image 110 , we can find 89.189 % of the lesion areas. In according to Eq 4.7 root mean square error value is: RMSE= 0.94.

4.8 Comparison Of All The Techniques

Comparison of the different unsupervised and supervised techniques according to false negative, false positive and true pixel numbers are given in the table 4.2.

Table 4.2 Comparison of the pattern recognition techniques in according to the pixel numbers for slices 100, 105 and 110.

Pattern Recognition Technique		FCA+PCA	K-Means +PCA	K-Medoid +PCA	ICA	ANN
Slice 100	Target pixels number	257	257	257	257	257
	True pixels in the result image	192	192	192	230	209
	False negative pixels	65	65	65	27	48
	False positive pixels	96	96	96	192	7
Slice 105	Target pixels number	180	180	180	180	180
	True pixels in the result image	94	94	94	62	143
	False negative pixels	86	86	86	118	37
	False positive pixels	1	1	1	0	2
Slice 110	Target pixels number	74	74	74	74	74
	True pixels in the result image	60	60	60	40	66
	False negative pixels	14	14	14	34	8
	False positive pixels	24	24	24	201	1

Another comparison critera is elapsed time. Comparison of the elapsed times for each algorithm are given in table 4.3.

Table 4.3 Comparison of the elapsed times, root mean square errors and result true pixels percentages of the different slices.

Pattern Recognition Technique	Percentage of the true pixels (%)			Elapsed Time (seconds)			RMSE		
	Slice 100	Slice 105	Slice 110	Slice 100	Slice 105	Slice 110	Slice 100	Slice 105	Slice 110
FCM+PCA	74.7	52.2	81.08	35.93	150.71	46.76	0.86	0.72	0.9
K-MEANS+PCA	74.7	52.2	81.08	8.75	12.50	14.97	0.86	0.72	0.9
K-MEDOID+PCA	74.7	52.2	81.08	5.35	13.16	5.75	0.86	0.72	0.9
ICA	89.49	34.44	54.05	0.41	0.40	0.34	0.94	0.58	0.73
ANN	81.32	79.44	89.18	0.14*	0.14*	0.21*	0.90	0.89	0.94

* These elapsed times are needed for simulation of the system. Training time for the system is 451.47 seconds with Intel Core2 CPU, 1.83 GHz and 1 GB RAM.

As can be seen from the tables, unsupervised clustering techniques are not as successful as supervised neural network techniques. The neural network system finds true pixels of the lesion areas with average % 83 success. The neural network system is trained with slice 100 and simulated with slice 100, 105 and 110 and multilayer feedforward backpropagation algorithm is used with 16 input, 16 hidden layers and 1 output layer.

We can compare these results different evaluation criterias as :

- Elapsed time.
- Percentage of the true pixels.
- False positive pixel numbers.
- False negative pixel numbers.
- Sensitivity
- Specificity

4.8.1 Comparison of Elapsed Times

Elapsed time is the time duration of the operating algorithm. Because of the artificial neural networks systems are needed to be train, it can be seen as the slowest technique. But, after training the system, simulation takes the shortest time duration between the all techniques. In clustering techniques, independent component analysis is the fastest and fuzzy c-means clustering with principle component analysis is the slowest technique. Although k-means and k-medoid algorithms are very similar to each other, k-medoid algorithm has better elapsed time score then k-means algorithm.

4.8.2 Comparison of Percentage of the True Pixels

This comparision is very much related to the location of the lesions in the processing image slice. For example, in slice 100 and 110, the lesions have contrast with CSF or white matter but in the slice 105, the lesions are distributed in the gray

matter and there is no very significant contrast between the neighbouring pixels. Thats why, we can see the lowest true pixel percentages on this image.

In general, again neural network system gives the most reliable results. If we consider component analysis techniques, fuzzy c-means, k-means and k-medoid clustering techniques with principle component analysis gives more accurate results then the independent component analysis.

4.8.3 Comparison of the False Positive Pixel Numbers

False positive pixels means that, there are some error pixels which are not supposed to be in the segmented image. Between these techniques, artificial neural network technique has less positive pixels then other clustering tehniques.

4.8.4 Comparison of the False Negative Pixel Numbers

False negative pixels means that there are some missing pixels in the result lesion area of the image. Artificial neural network system gives less false negative pixels then the others. Fuzzy C-Means, K-Means and K-Medoid clustering techniques are same false negative pixel numbers. Although independent component analysis has the best score in the false positive pixel number, according to this criteria, it has more missing pixels then the others. There is an erosion between lesion areas and healty brain tissues in the independent component analysis.

4.8.5 Sensitivity

The calculation of sensitivity is:

$$Sensitivity = \frac{TP}{TP + FN}$$

TP=True Pixels

FN=False Negative Pixels

4.8.5.1 Sensitivity of clustering methods

Fuzzy C-Means, K-Means and K-Medoid Clustering Methods gives very close results:

For slice 100:

$$Sensitivity = \frac{192}{192+65} = 0.74$$

For slice 105:

$$Sensitivity = \frac{94}{94+86} = 0.52$$

For slice 110:

$$Sensitivity = \frac{60}{60+14} = 0.81$$

Average sensitivity is 0.69

4.8.5.2 Sensitivity of independent component analysis method

The sensitivity calculation of the independent component analysis method is below:

For slice 100:

$$Sensitivity = \frac{230}{230+27} = 0.89$$

For slice 105:

$$Sensitivity = \frac{62}{62+118} = 0.34$$

For slice 110:

$$Sensitivity = \frac{40}{40+34} = 0.54$$

Average sensitivity is 0.59

4.8.5.3 Sensitivity of multilayer backpropagation neural network method

The sensitivity calculation of multilayer backpropagation neural network method is below:

For slice 100:

$$Sensitivity = \frac{209}{209+48} = 0.81$$

For slice 105:

$$Sensitivity = \frac{143}{143+37} = 0.79$$

For slice 110:

$$Sensitivity = \frac{66}{66+8} = 0.89$$

Average sensitivity is 0.83

4.8.6 Specificity

Calculation of the specificity is below:

$$Specificity = \frac{TN}{TN+FP}$$

TN=True Negatives.

FP=False Positives.

4.8.6.1 Specificity of clustering methods

Fuzzy C-Means, K-Means and K-Medoid Clustering Methods gives very close results:

For slice 100:

$$Specificity = \frac{39020}{39020+96} = 0.998$$

For slice 105:

$$Specificity = \frac{39097}{39097+1} = 0.999$$

For slice 110:

$$Specificity = \frac{39203}{39203+24} = 0.999$$

Average specificity is 0.998

4.8.6.2 Specificity of independent component analysis method

The specificity calculation of the independent component analysis method is below:

For slice 100:

$$Specificity = \frac{39020}{39020+192} = 0.995$$

For slice 105:

$$Specificity = \frac{39097}{39097+0} = 1$$

For slice 110:

$$Specificity = \frac{39203}{39203+201} = 0.994$$

Average specificity is 0.996

4.8.6.3 Specificity of multilayer backpropagation neural network method

The specificity calculation of multilayer backpropagation neural network method is below:

For slice 100:

$$Specificity = \frac{39020}{39020+7} = 0.999$$

For slice 105:

$$Specificity = \frac{39097}{39097+2} = 0.999$$

For slice 110:

$$Specificity = \frac{39203}{39203+1} = 0.999$$

Average specificity is 0.999

4.9 Result

According to comparison of the sensitivity and specificity criterias, multilayer backpropagation neural network algorithm has the best sensitivity and specificity between all techniques.

Clustering techniques with principle component analysis, are better then the independent component analysis's sensitivity and specificity.

4.10 GUI

In order to compare different segmentation methods, a segmentation GUI has been developed. MATLAB is necessary to execute the segmentation GUI. Figure 4.63 shows this GUI.

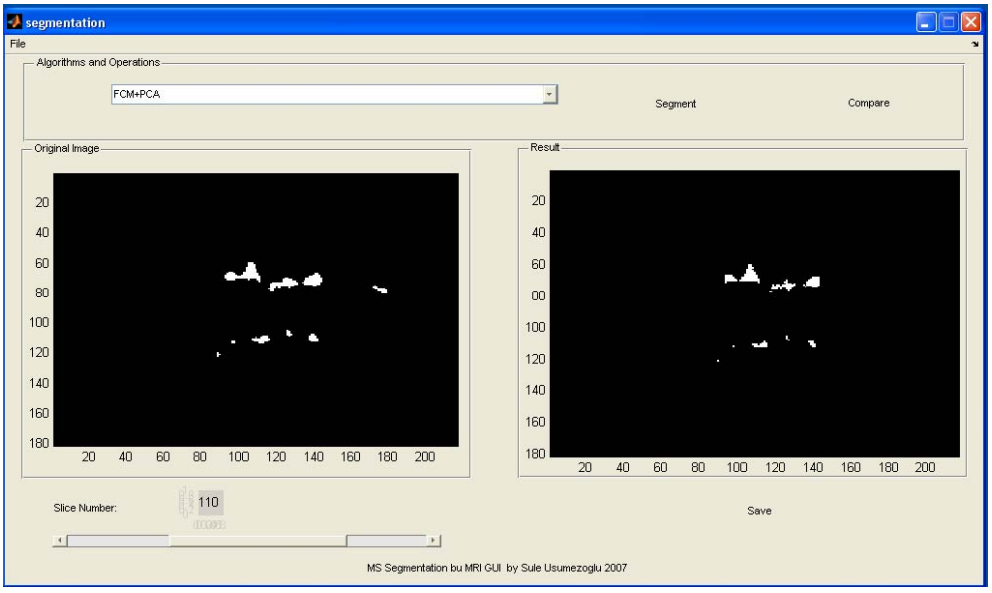


Figure 4.63 Segmentation GUI.

The first part of the GUI is to select the slice number and segmentation type. Figure 4.64 shows the algorithms that can be selected. Slice number can be selected from the floading bar on the left bottom of the GUI. After selecting the slice number, segmentation type is selected from the window. Segment button starts the

segmentation for selected algorithm. After the segmentation process, the left side of the GUI shows the original “gold” MS lesions image and right side shows the segmented image. We can see the differances from this window.

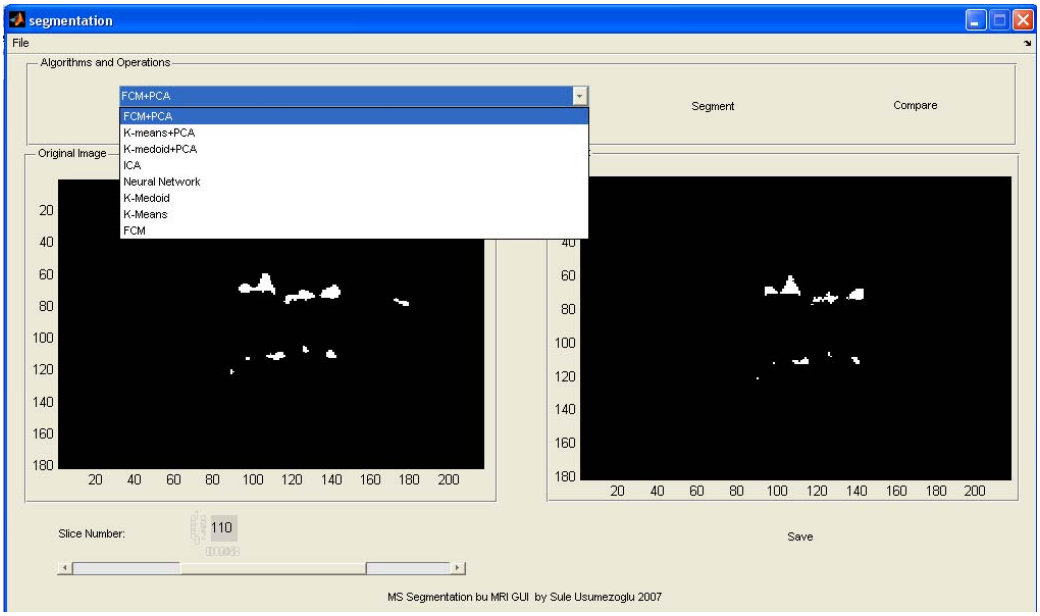
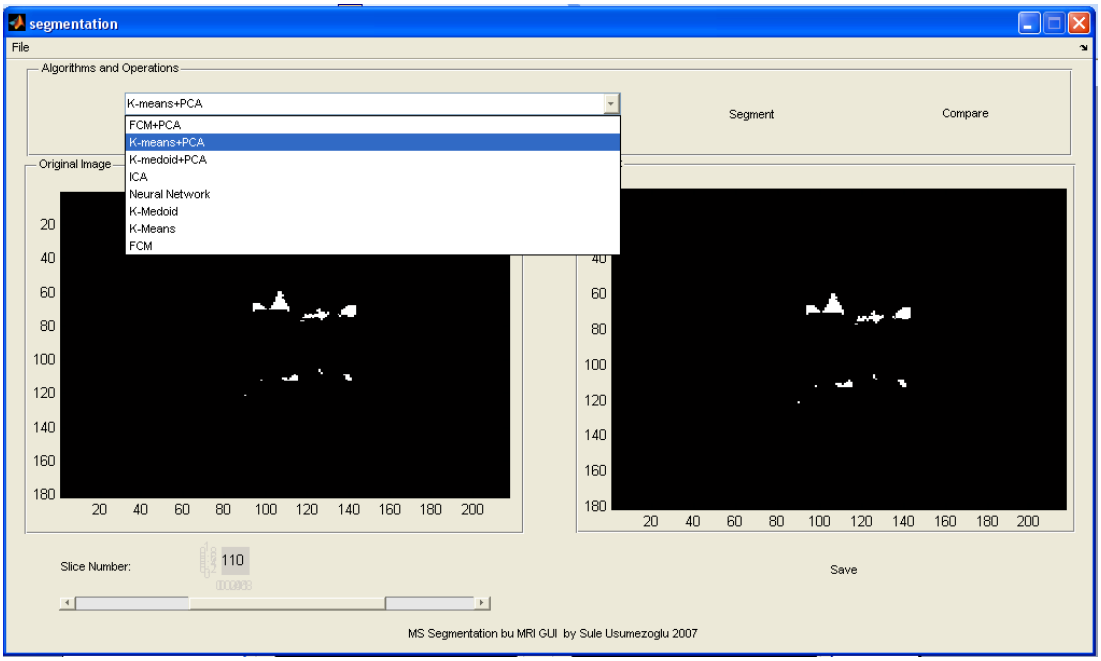


Figure 4.64 Selection of the segmentation types and slice number.

After saving the segmentated image, we can choose another algorithm and compare them with compare button. Figure 4.65 shows the comparison of the Fuzzy C-Means+PCA and K-Means+PCA algorithms. This GUI can be used to compare the different segmentation algorithms as Independent Component Analysis and Multilayer Back Propagation Neural Network Algorithm.



4.65 Comparision of the segmentation methods by using segmentation GUI.

CHAPTER FIVE
CONCLUSION

In this thesis, different segmentation methods are applied to the efficient automatic segmentation of Multiple Sclerosis (MS) for obtaining lesions from T1, T2, PD weighted brain MR images. These systems are based on pixel classifications.

Multilayer backpropagation neural networks are considered as supervised segmentation techniques and clustering algorithms are considered as unsupervised techniques. Fuzzy C-Means, K-Means and K-medoid clustering techniques with Principle Component Analysis, Independent Component Analysis and Multi Layer Back Propagation Neural Network algorithms are compared and results are evaluated.

Results are compared based on different evaluation criteria as elapsed time, percentage of the true pixels, false positive pixel numbers, false negative pixel numbers, sensitivity and specificity of the result images.

Because of the inhomogenities of the MRI images, some lesions are overlapping with the gray matter and this situation causes some false negative pixels on the result images.

The selection of the threshold value that specifies lesions areas, is important for to find accurate lesion areas . Threshold value is selected based on the true pixels-false pixels graphics. The same threshold value is applied to all segmentation methods.

Sensitivity of the segmented images is confirmed with the "gold" MS lesions images which belongs to the related MRI slices. This comparison has shown that, the segmentation success is very related to the location of the lesions in the processing image slice. For example, in the slice 100 and 110, the lesions have contrast with CSF or white matter but in the slice 105, lesions are distributed in the

gray matter and there is no very significant contrast between the neighbouring pixels. That' s why, on this image.we can see the lowest true pixel percentages

If we compare unsupervised hard and soft partition methods, we can see that the soft partition method gives more successful results than hard partition methods and have more elapsed time. The reason for this, is that in the hard partition methods a pixel can belong to only one cluster but in the soft partition methods a pixel can belong to more than one cluster and MRI images are better treated this type of structure.

In general, supervised artificial neural network method has given the best, most reliable and accurate results. Despite the fact that is requires training time, simulations afterwards are faster then any other technique.

5.1 Future Work

The results in this thesis are very encouraging about segmentation of MS lesions from MRI images. The combination of the supervised and unsupervised methods can give more accurate results and clustering techiques can be used to obtain training data where there is no available data for train of the supervised algorithms.

An automated decision algorithm can be develop in order to choose the most suitable clustered images from the clustering algorithms with image comparison techniques.

The combination of supervised and unsupervised techniques can be proposed as hybrid method and these algorithms can be implement in 3D images.

Preprocessing of MRI images with bias filter correction or other image processing techniques can be improved.

In this thesis, multilayer backpropagation algorithm as neural approach is investigated. Other neural network methods can also be applied like learning quantization algorithms or self organizing maps for segmentation of brain tissues with morphological postprocessing.

REFERENCES

Balasko, B., Aboyni J., Feil, B. (2006). Fuzzy clustering and data analysis toolbox. *Proceedings of 7th International Symposium of Hungarian Researchers on Computational Intelligence*.(April 2007). http://www.mathworks.com/matlabcentral/fileexchange/loadFile.do?objectId=748

Behan, P., & Chaudhuri, A. (2002). The pathogenesis of multiple sclerosis revisited. *J R Coll Physicians Edinb.* 32: 244–265.

Bezdek, J.C., & Dunn, J.C. (1975). Optimal Fuzzy partitions: A herustic for estimating the parameters in a mixture of normal distributions. *IEEE transactions on computers*, pages 835-838.

Calabresi, P.A. (2004). Diagnosis and management of multiple sclerosis. *American Familiy Physician.*

Cichocki, A., Amari, S., Siwek, K., & Tanaka, T. Icalab Toolboxes. (April, 2007). http://www.bsp.brain.riken.jp/ICALAB.

Cocosco, C. A., Kollokian, V., Kwan, R. K.-S., Evans, A.C. (1997). Brainweb: Online Interface to a 3D MRI Simulated Brain Database. *NeuroImage,* vol.5, no.4, part 2/4, S425. Proceedings of 3th International Conference on Functional Mapping of the Human Brain, Copenhagen, May 1997. BrainWeb:Simulated Brain Database. (April, 2007). http://www.bic.mni.mcgill.ca/brainweb/.

Confavreux, C., Suissa, S., Saddier, P., Bourdès, V., & Vukusic, S. (2001). Vaccinations and the risk of relapse in multiple sclerosis. Vaccines in Multiple Sclerosis Study Group. *N. Engl. J. Med.* 344 (5): 319-26.

Cuevas J. E., Zaldivar N. D., Rojas R. (2004). Fuzzy segmentation in image processing. *XXVI International Congress on Electrical Engineering, Electro* 2004.

Dangond, F. (2005). Multiple sclerosis. *eMedicine Neurology*. November 2004.

Duda, R., O., Hart, P. E., & Stork D. G. (2001). *Pattern classification*. Second edition, Wiley Interscience Publication.

Gronseth, G., Ashman, E., (2000). Practice parameter: the usefulness of evoked potentials in identifying clinically silent lesions in patients with suspected multiple sclerosis (an evidence-based review): Report of the Quality Standards Subcommittee of the American Academy of Neurology. *Neurology* 2000 May 9;54(9):1720–5.

Gustafson, D. E., & Kessel, W. C. (1979). Fuzzy clustering with fuzzy covariance matrix. *In proceedings of the IEEE CDC,* San Diego, pages 761-766.1979

Ham, F. M., & Kostanic I. (2001). *Principles of neurocomputing for science and engineering.* McGrrow-Hill International Edition.

Hornac, J. P.(2007). *The basis of MRI*. (April, 2007). Copyright © 1996-2007. http://www.cis.rit.edu/htbooks/mri.

http://en.wikipedia.org/wiki/Multiple_sclerosis#_note-42 (June, 2007).

Jiri, J. (2006). *Medical image processing, reconstruction and restoration concepts and methods.* Taylor&Francis Group.

Jongen, P. (2006). Psychiatric onset of multiple sclerosis. *J Neurol Sci* 245 (1–2): 59–62.

Kurtzke, J.F. (1983). Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). 1983 Nov;33(11):1444-52.

Lucchinetti, C., Parisi, W., J. Scherhauer, B. Rodriguez, M. Lassmann, H.(2000). Heterogeneity of multiple sclerosis lesions: implications for the pathogenesis of demyelination . *Ann Neurol*, 2000; 47(6):707-17.

Marrie, R. (2004). Environmental risk factors in multiple sclerosis aetiology. *Lancet Neurol*. 2004 Dec;3(12):709-18.

Mbah, H. O. (2006). *Construction of magnetic resonance images.* African Institute for Mathematical Sciences, Cape Town.

McDonald, W., Compston, A., Edan, G., Goodkin, D., Hartung, H.,Lublin, F., McFarland, H., Paty, D., Polman, C., Reingold, S., Sandberg-Wollheim, M., Sibley, W., Thompson, A., van den Noort, S., Weinshenker, B., & Wolinsky, J.(2001). Recommended diagnostic criteria for multiple sclerosis: guidelines from the International Panel on the diagnosis of multiple sclerosis. *Ann Neurol* 2001 Jul;50(1):121-7

Narayana, P. A., Sajja, B. R., Datta, S., & He, R. (2004). A unified approach for lesion segmentation on mri of mutiple sclerosis*. Proceedings of the 26th Annual International Conference of the IEEE EMBS,* San Francisco, CA.

Navarro, S., Mondéjar-Marín, B., Pedrosa-Guerrero, A., Pérez-Molina, I., Garrido-Robres, J., & Alvarez-Tejerina, A. (2005). Aphasia and parietal syndrome as the presenting symptoms of a demyelinating disease with pseudotumoral lesions. *Rev Neurol 41* (10): 601-3.

Özkurt, A., & Özkurt, N. (2007*).* MR görüntülerinin bağımsız bileşen analizi yöntemiyle güdümsüz bölütlenmesi. *IEEE SIU'07 Kurultayı*, Eskişehir.

Parra, C. A., Iftekharuddin K., & Kozma R. (2003). Automated brain data segmentation and pattern recognition using ANN. *Computational Intelligence, Robotics and Autonomous Systems* (CIRAS 03).

Pascual, A., Martínez-Bisbal, M., Boscá I, *et al* (2007). Axonal loss is progressive and partly dissociated from lesion load in early multiple sclerosis. *Neurology* 69 (1): 63-7.

Paty, D., Studney, D., Redekop, K., & Lublin, F. (1994) MS COSTAR: a computerized patient record adapted for clinical research purposes. *Ann Neurol* 1994;36 Suppl:S134-5

Pu, M. (2001). *Automatic segmentation of multiple sclerosis lesions in magnetic resonance brain images.* Dalhousie University, Halifax, Nova Scotia.

Ren, H., & Chang, C.I. (2000*).* A generalized orthogonal subspace projection approach to unsupervised multispectral image classification. *IEEE Trans. on Geoscience and Remote Sensing*, 38(6), 2515-2528.

Rentzos, M., Nikolaou, C., Anagnostouli, M., Rombos, A., Tsakanikas, K., Economou, M., Dimitrakopoulos, A., Karouli, M., & Vassilopoulos, D. (2006). Serum uric acid and multiple sclerosis. *Clinical neurology and neurosurgery* 108 (6): 527-31.

Rudick, R., & Whitaker, J., Cerebrospinal fluid tests for multiple sclerosis*. In Scheinberg, P (Ed). Neurology/neurosurgery update series,* Vol. 7, CPEC. Princeton, NJ 1987

Rothwell, P., & Charlton, D. (1998). High incidence and prevalence of multiple sclerosis in south east Scotland: Evidence of a genetic predisposition. *J. Neurol. Neurosurg. Psychiatr.* 64 (6): 730-5.

Sadovnick, A., Ebers, G., Dyment, D., Risch, N.(1996). Evidence for genetic basis of multiple sclerosis. *The Canadian Collaborative Study Group.* Lancet 1996; 347:1728.

Tipping, M. E., & Bishop, C. M. (1999). Mixtures of probabilistic principle component analysis. *Neural Computation*, 11: 443-482.

Yen, J., & Langhari, R. (2000). *Fuzzy logic, intelligence, control and information.* Prentice Hall, New York.

Weinshenker, B. (2005). Western vs optic-spinal MS: Two diseases, one treatment?. *Neurology* 64 (4): 594-5.

Zheng, L., & Robert, L. (2006). Phase congruence measurement for image similarity assessment. *Science Direct, Pattern Recognition Letters*, Elseiver.

APPENDIX

MATLAB codes:

i.) Code for fuzzy c-means clustering with principle component analysis:

```
close all
clear all

fid1 = fopen('t1.rawb', 'r');
c1 = fread(fid1,'uint8');
fclose(fid1);

fid2 = fopen('t2.rawb', 'r');
c2 = fread(fid2,'uint8');
fclose(fid2);

fid3 = fopen('pd.rawb', 'r');
c3 = fread(fid3,'uint8');
fclose(fid3);

fid4 = fopen('hedef.rawb', 'r');
c4 = fread(fid4,'uint8');
fclose(fid4);

 A1=reshape(c1,181,217,181);
A2=reshape(c2,181,217,181);
A3=reshape(c3,181,217,181);
A4=reshape(c4,181,217,181);


K=110;
cc1=A1(:,:,K);
cc1=double(cc1);
cc2=A2(:,:,K);
cc2=double(cc2);
cc3=A3(:,:,K);
cc3=double(cc3);
cc4=A4(:,:,K);
cc4=double(cc4);


b=zeros(size(cc4));

b(find(cc4>=80))=255; %test image
b(find(cc4<80))=0;


[m n]=size(cc1);
s11=reshape(cc1,1,m*n);
s22=reshape(cc2,1,m*n);
s33=reshape(cc3,1,m*n);
s44=reshape(cc4,1,m*n);
 s=zeros(24,m*n);
```

```
s(1,:)=s11;
s(2,:)=s22;
s(3,:)=s33;
s(4,:)=s11.*s22;
s(5,:)=s11.*s33;
s(6,:)=s22.*s33;
s(7,:)=sqrt(s11.*s22);
s(8,:)=sqrt(s11.*s33);
s(9,:)=sqrt(s22.*s33);
s(10,:)=s11.^2;
s(11,:)=s22.^2;
s(12,:)=s33.^2;
s(13,:)=sqrt(s11);
s(14,:)=sqrt(s22);
s(15,:)=sqrt(s33);
s(16,:)=s11.*(255-s22);
s(17,:)=s22.*(255-s33);
s(18,:)=s33.*(255-s11);
s(19,:)=s22.*(255-s11);
s(20,:)=s33.*(255-s22);
s(21,:)=s11.*(255-s33);
s(22,:)=(255-s11);
s(23,:)=(255-s22);
s(24,:)=(255-s33);


data.X=s';



data = clust_normalize(data,'range')
subplot(2,1,1),plot(data.X(:,1),data.X(:,2),'x')
data = clust_normalize(data,'var')
subplot(2,1,2),plot(data.X(:,1),data.X(:,2),'+')


tic
param.c = 6;
param.m = 24;
result = FCMclust(data,param)


param.q = 6;
result = PCA(data,param,result);

toc
%
Y=result.proj.P;
%
y=Y';
%

[k,l]=size(y);
```

```
figure(2),
t=1;
for i=1:6
   subplot(3,2,t),imagesc(reshape(y(i,:),m,n)); colormap(gray);
   title(strcat(num2str(i),'. FCM+PCA'));colormap(gray)
   t=t+1;
end


%%%% find region




im=[];
ics=[];
for i=1:k
   im=reshape(y(i,:),m,n);
   im=im-min(min(im));
   im=im*(255/max(max(im)));
   ics(:,:,i)=im;

end

result=ics(:,:,j(6));
result=round(result);



a=[0:255]/256;
aa=[a' a' a'];

figure,
subplot(121), imagesc(im),colormap(gray),title('(a) clustered image');
subplot(122),imhist(im,aa),title('(b) histogram');


 result1=zeros(size(im));

 result1(find(im>=40))=255;
result1(find(im<40))=0;

result2=zeros(size(im));
result2(find(result1==0))=255;

figure(4),imagesc(result2),colormap(gray);

 figure(5),
 subplot(221),imagesc(cc1),colormap(gray),title ('(a) Original Image');
 subplot(222),imagesc(b),colormap(gray),title ('(b) Target Image');
 subplot(223),imagesc(result),colormap(gray),title ('(c) FCM+PCA');
 subplot(224),imagesc(result2),colormap(gray),title ('(d) Result Image');
```

```
testim=zeros(m,n);
% b is the target image
% result is the segmented image
for i=1:m;
   for j=1:n;
      if b(i,j)==result2(i,j);
         testim(i,j)=0;
         else testim(i,j)=255;
      end

   end
end

% testim iamge gives difference between target
% and result images

figure(6), imagesc(b),colormap(gray), title('target');
figure(7), imagesc(result2),colormap(gray), title('result');

figure(8), imagesc(testim),colormap(gray), title('difference between real target and result');



%calculate error
ms=length(find(b==255))%target pixel number
tp=length(find((b==255)&(result2==255)))

w1=ms-tp %false negative pixels
ts=length(find(testim==255))

rs=length(find(result2==255))
w2=rs-tp %false positive pixels


%percent of true pixels
tt=(tp/ms)*100

%%%%%%%%%%%%%%%%%%%%%
%%%%calculation of RMSE
%%%%%%%%%%%%%%%%%%%%%

rmse=sqrt(tp/ms)
```

ii.) Code for K-means clustering with principle component analysis:

```
close all
clear all


fid1 = fopen('t1.rawb', 'r');
c1 = fread(fid1,'uint8');
fclose(fid1);
```

```
fid2 = fopen('t2.rawb', 'r');
c2 = fread(fid2,'uint8');
fclose(fid2);

fid3 = fopen('pd.rawb', 'r');
c3 = fread(fid3,'uint8');
fclose(fid3);

fid4 = fopen('hedef.rawb', 'r');
c4 = fread(fid4,'uint8');
fclose(fid4);

A1=reshape(c1,181,217,181);
A2=reshape(c2,181,217,181);
A3=reshape(c3,181,217,181);
A4=reshape(c4,181,217,181);


K=110;
cc1=A1(:,:,K);
cc1=double(cc1);
cc2=A2(:,:,K);
cc2=double(cc2);
cc3=A3(:,:,K);
cc3=double(cc3);
cc4=A4(:,:,K);
cc4=double(cc4);

b=zeros(size(cc4));

b(find(cc4>=80))=255; %test image
b(find(cc4<80))=0;


figure(1),imagesc(b),colormap(gray),title('target image');




[m n]=size(cc1);
s11=reshape(cc1,1,m*n);
s22=reshape(cc2,1,m*n);
s33=reshape(cc3,1,m*n);
s44=reshape(cc4,1,m*n);

s=zeros(24,m*n);



s(1,:)=s11;
s(2,:)=s22;
s(3,:)=s33;
s(4,:)=s11.*s22;
s(5,:)=s11.*s33;
```

```
s(6,:)=s22.*s33;
s(7,:)=sqrt(s11.*s22);
s(8,:)=sqrt(s11.*s33);
s(9,:)=sqrt(s22.*s33);
s(10,:)=s11.^2;
s(11,:)=s22.^2;
s(12,:)=s33.^2;
s(13,:)=sqrt(s11);
s(14,:)=sqrt(s22);
s(15,:)=sqrt(s33);
s(16,:)=s11.*(255-s22);
s(17,:)=s22.*(255-s33);
s(18,:)=s33.*(255-s11);
s(19,:)=s22.*(255-s11);
s(20,:)=s33.*(255-s22);
s(21,:)=s11.*(255-s33);
s(22,:)=(255-s11);
s(23,:)=(255-s22);
s(24,:)=(255-s33);




data.X=s';

data = clust_normalize(data,'range')
subplot(3,1,2),plot(data.X(:,1),data.X(:,2),'x')
data = clust_normalize(data,'var')
subplot(3,1,3),plot(data.X(:,1),data.X(:,2),'+')


tic
param.vis=1;
param.c=6;
param.m=4;
param.val=2;


result=Keymeans(data,param);


param.q = 6;
result = PCA(data,param,result);
toc
H=result.data.d;


%
Y=result.proj.P;
%
y=Y';
%

[k,l]=size(y);
```

```
figure,
t=1;
for i=1:6
   subplot(3,2,t),imagesc(reshape(y(i,:),m,n)); colormap(gray);
   title(strcat(num2str(i),'. KMEANS+PCA'));colormap(gray)
   t=t+1;
end


%%%% find region




im=[];
ics=[];
for i=1:k
   im=reshape(y(i,:),m,n);
   im=im-min(min(im));
   im=im*(255/max(max(im)));
   ics(:,:,i)=im;

end
result=ics(:,:,j(4));
result=round(result);

a=[0:255]/256;
aa=[a' a' a'];

 figure,
 subplot(121),imagesc(result),colormap(gray),title('(a) K-Means Clustered image');
 subplot(122),imhist(result,aa),title('(b) Histogram');



 result1=result;
 result1(find(result1>=40))=255;
 result1(find(result1<40))=0;

 result2=zeros(size(result1));
 result2(find(result1==0))=255;


  figure,
 imagesc(result2),colormap(gray);



 figure,
  subplot(221),imagesc(cc1),colormap(gray),title('(a) original T1 image');
```

```
  subplot(222),imagesc(b),colormap(gray),title('(b) target  image');
  subplot(223),imagesc(result),colormap(gray),title('(c) kmeans clustered image');
  subplot(224),imagesc(result2),colormap(gray),title('(d) result image');

%calculate error
ms=length(find(b==255))%target pixel number
tp=length(find((b==255)&(result2==255)))
% h %true pixel number of the result image
w1=ms-tp %false negative pixels
ts=length(find(testim==255))

rs=length(find(result2==255))
w2=rs-tp %false positive pixels


%percent of true pixels
tt=(tp/ms)*100

%%%%%%%%%%%%%%%%%%%%
%%%%calculation of RMSE
%%%%%%%%%%%%%%%%%%%%

rmse=sqrt(tp/ms)
```

iii.) Code for K-medoid clustering with principle component analysis:

```
close all
clear all


fid1 = fopen('t1.rawb', 'r');
c1 = fread(fid1,'uint8');
fclose(fid1);

fid2 = fopen('t2.rawb', 'r');
c2 = fread(fid2,'uint8');
fclose(fid2);

fid3 = fopen('pd.rawb', 'r');
c3 = fread(fid3,'uint8');
fclose(fid3);

fid4 = fopen('hedef.rawb', 'r');
c4 = fread(fid4,'uint8');
fclose(fid4);

A1=reshape(c1,181,217,181);
A2=reshape(c2,181,217,181);
A3=reshape(c3,181,217,181);
A4=reshape(c4,181,217,181);

K=110;
cc1=A1(:,:,K);
```

```
cc2=A2(:,:,K);

cc3=A3(:,:,K);

cc4=A4(:,:,K);

b=zeros(size(cc4));

b(find(cc4>=80))=255; %test image
b(find(cc4<80))=0;

[m n]=size(cc4);

s11=reshape(double(cc1),1,m*n);
s22=reshape(double(cc2),1,m*n);
s33=reshape(double(cc3),1,m*n);
s44=reshape(double(cc4),1,m*n);

s=zeros(24,m*n);



s(1,:)=s11;
s(2,:)=s22;
s(3,:)=s33;
s(4,:)=s11.*s22;
s(5,:)=s11.*s33;
s(6,:)=s22.*s33;
s(7,:)=sqrt(s11.*s22);
s(8,:)=sqrt(s11.*s33);
s(9,:)=sqrt(s22.*s33);
s(10,:)=s11.^2;
s(11,:)=s22.^2;
s(12,:)=s33.^2;
s(13,:)=sqrt(s11);
s(14,:)=sqrt(s22);
s(15,:)=sqrt(s33);
s(16,:)=s11.*(255-s22);
s(17,:)=s22.*(255-s33);
s(18,:)=s33.*(255-s11);
s(19,:)=s22.*(255-s11);
s(20,:)=s33.*(255-s22);
s(21,:)=s11.*(255-s33);
s(22,:)=(255-s11);
s(23,:)=(255-s22);
s(24,:)=(255-s33);

 data.X=s';

data = clust_normalize(data,'range')
subplot(3,1,2),plot(data.X(:,1),data.X(:,2),'x')
data = clust_normalize(data,'var')
subplot(3,1,3),plot(data.X(:,1),data.X(:,2),'+')
```

```
tic
param.vis=1;
param.c=6;
param.m=24;
param.val=2;
result=Kmedoid(data,param);

param.q = 6;
result = PCA(data,param,result);
toc


a=[0:255]/256;
aa=[a' a' a'];


Y=result.proj.P;
%
y=Y';
%


figure,
[k,l]=size(y);
for i=1:k
   no=round(k/6);
   subplot(no,6,i),imagesc(reshape(y(i,:),m,n)); colormap(gray);
end

figure,
t=1;
for i=1:6
   subplot(3,2,t),imagesc(reshape(y(i,:),m,n)); colormap(gray);
   title(strcat(num2str(i),'.Kmedoid+pca'));colormap(gray)
   t=t+1;
end


%%%% find region



im=[];
ics=[];
for i=1:k
   im=reshape(y(i,:),m,n);
   im=im-min(min(im));
   im=im*(255/max(max(im)));
   ics(:,:,i)=im;

end
result=ics(:,:,j(4));
result=round(result);
```

```
figure,
subplot(121),imagesc(result),colormap(gray),title('(a) K-Medoid and PCA Clustered Image');
subplot(122),imhist(result,aa),title('(b) Histogram');

result1=result;
result1(find(result1>=40))=255;
result1(find(result1<40))=0;

result2=zeros(size(result1));
result2(find(result1==0))=255;


figure,
imagesc(result2),colormap(gray);

figure,
subplot(221),imagesc(cc1),colormap(gray),title('(a) original T1 image');
subplot(222),imagesc(b),colormap(gray),title('(b) target  image');
subplot(223),imagesc(result),colormap(gray),title('(c) kmedoid clustered image');
subplot(224),imagesc(result2),colormap(gray),title('(d) result image');

testim=zeros(m,n);

for i=1:m;
   for j=1:n;
      if b(i,j)==result2(i,j);
         testim(i,j)=0;
         else testim(i,j)=255;
      end

   end
end




figure, imagesc(b),colormap(gray), title('target');
figure, imagesc(result2),colormap(gray), title('result');

figure, imagesc(testim),colormap(gray), title('difference between real target and result');

%calculate error
ms=length(find(b==255))%target pixel number
tp=length(find((b==255)&(result2==255)))
% h %true pixel number of the result image
w1=ms-tp %false negative pixels
ts=length(find(testim==255))

rs=length(find(result2==255))
w2=rs-tp %false positive pixels
```

```
%percent of true pixels
tt=(tp/ms)*100

%%%%%%%%%%%%%%%%%%%%%%
%%%%calculation of RMSE
%%%%%%%%%%%%%%%%%%%%%%

rmse=sqrt(tp/ms)
```

iv.) Code for independent component analysis:

```
clear all
close all

fid1 = fopen('t1.rawb', 'r');
c1 = fread(fid1,'uint8');
fclose(fid1);
fid2 = fopen('t2.rawb', 'r');
c2 = fread(fid2,'uint8');
fclose(fid2);
fid3 = fopen('pd.rawb', 'r');
c3 = fread(fid3,'uint8');
fclose(fid3);
fid4 = fopen('hedef.rawb', 'r');
c4 = fread(fid4,'uint8');
fclose(fid4);

A1=reshape(c1,181,217,181);
A2=reshape(c2,181,217,181);
A3=reshape(c3,181,217,181);
A4=reshape(c4,181,217,181);


m=181;
n=217;

K=110; %slice number

cc1=A1(:,:,K);
cc2=A2(:,:,K);
cc3=A3(:,:,K);
cc4=A4(:,:,K);

cc1(find(cc1<25))=0;
cc2(find(cc2<25))=0;
cc3(find(cc3<25))=0;
cc4=double(cc4);

b=zeros(size(cc4));
b(find(cc4>=80))=255; %test image
b(find(cc4<80))=0;
```

```
s11=reshape(double(cc1),1,m*n);
s22=reshape(double(cc2),1,m*n);
s33=reshape(double(cc3),1,m*n);

s(1,:)=s11;
s(2,:)=s22;
s(3,:)=s33;
s(4,:)=s11.*s22;
s(5,:)=s11.*s33;
s(6,:)=s22.*s33;
s(7,:)=sqrt(s11.*s22);
s(8,:)=sqrt(s11.*s33);
s(9,:)=sqrt(s22.*s33);
s(10,:)=s11.^2;
s(11,:)=s22.^2;
s(12,:)=s33.^2;
s(13,:)=sqrt(s11);
s(14,:)=sqrt(s22);
s(15,:)=sqrt(s33);
s(16,:)=s11.*(255-s22);
s(17,:)=s22.*(255-s33);
s(18,:)=s33.*(255-s11);
s(19,:)=s22.*(255-s11);
s(20,:)=s33.*(255-s22);
s(21,:)=s11.*(255-s33);
s(22,:)=(255-s11);
s(23,:)=(255-s22);
s(24,:)=(255-s33);
tic
W=amuse(s);
y=W*s;
toc


figure(1),
[k,l]=size(y);
for i=1:k
   no=round(k/6);
   subplot(no,6,i),imagesc(reshape(y(i,:),m,n)); colormap(gray);
   if i==1 title('Independent components'); end
end

% figure, imagesc(reshape(y(7,:),m,n)),colormap(gray)


im=[];
ics=[];
[k,l]=size(y);
for i=1:k
   im=reshape(y(i,:),m,n);
   im=im-min(min(im));
   im=im*(255/max(max(im)));
   ics(:,:,i)=im;

end
```

```
result=ics(:,:,j(5));

a=[0:255]/256;
aa=[a' a' a'];

figure(2),
subplot(121), imagesc(result),colormap(gray),title('(a) ICA clustered image');
subplot(122),imhist(result,aa),title('(b) histogram');


 result1=zeros(size(result));




 result1(find(result>=185))=255;
 result1(find(result<185))=0;



 figure,imagesc(result1),colormap(gray);



 figure(4),
  subplot(221),imagesc(cc1),colormap(gray),title ('(a) Original Image');
  subplot(222),imagesc(b),colormap(gray),title ('(b) Target Image');
  subplot(223),imagesc(result),colormap(gray),title ('(c) ICA clustered image');
  subplot(224),imagesc(result1),colormap(gray),title ('(d) Result Image');



  result2=result1;


  testim=zeros(m,n);


for i=1:m;
   for j=1:n;
      if b(i,j)==result2(i,j);
         testim(i,j)=0;
         else testim(i,j)=255;
      end

   end
end




figure(5), imagesc(b),colormap(gray), title('target');
```

```
figure(6), imagesc(result2),colormap(gray), title('result');

figure(7), imagesc(testim),colormap(gray), title('difference between real target and result');



%calculate error
ms=length(find(b==255))%target pixel number
tp=length(find((b==255)&(result1==255)))
% h %true pixel number of the result image
w1=ms-tp %false negative pixels
ts=length(find(testim==255))

rs=length(find(result1==255))
w2=rs-tp %false positive pixels


%percent of true pixels
tt=(tp/ms)*100

%%%%%%%%%%%%%%%%%%%%
%%%%calculation of RMSE
%%%%%%%%%%%%%%%%%%%%

rmse=sqrt(tp/ms)
%%%%%%%%%%%%%%%%%%%%
```

v.) Code for segmentation with neural networks:

```
close all
clear all


fid1 = fopen('t1.rawb', 'r');
c1 = fread(fid1,'uint8');
fclose(fid1);

fid2 = fopen('t2.rawb', 'r');
c2 = fread(fid2,'uint8');
fclose(fid2);

fid3 = fopen('pd.rawb', 'r');
c3 = fread(fid3,'uint8');
fclose(fid3);

fid4 = fopen('hedef.rawb', 'r');
c4 = fread(fid4,'uint8');
fclose(fid4);

A1=reshape(c1,181,217,181);
A2=reshape(c2,181,217,181);
A3=reshape(c3,181,217,181);
A4=reshape(c4,181,217,181);
```

```
K=100;


cc1=A1(:,:,K);

cc2=A2(:,:,K);

cc3=A3(:,:,K);

cc4=A4(:,:,K);
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%


b=zeros(size(cc4));

b(find(cc4>=80))=255; %test image
b(find(cc4<80))=0;

%figure,imagesc(b),colormap(gray),title('target image');



%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
[m,n]=size(cc1);




cc1(find(cc1<25))=0;
cc2(find(cc2<25))=0;
cc3(find(cc3<25))=0;
cc4(find(cc4<25))=0;


%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%




s11=reshape(double(cc1),1,m*n);
s22=reshape(double(cc2),1,m*n);
s33=reshape(double(cc3),1,m*n);
s44=reshape(double(cc4),1,m*n);



s=zeros(3,m*n);
```

```
s(1,:)=s11;
s(2,:)=s22;
s(3,:)=s33;




inp=s;

out=s44;
tic
network=newff(minmax(inp),[16,1],{'tansig','purelin'});
network=init(network);

network.trainParam.epochs = 400;
network=train(network,inp,out);
toc

y=sim(network,inp);

% Layer1_Weights=network.iw{1};
% Layer1_Bias=network.b{1};
% Layer2_Weights=network.lw{2};
% Layer2_Bias=network.b{2};
save Layer1_Weights ;
save Layer1_Bias;
save Layer2_Weights;
save Layer2_Bias;
Actual_Desired=[y' out'];
Actual_Desired;

%




[k,l]=size(y);




%%%% find region



im=[];
```

```
ics=[];

   im=reshape(y(1,:),m,n);
   im=im-min(min(im));
   im=im*(255/max(max(im)));
   ics(:,:)=im;

result=ics(:,:,j(1));
result=round(result);

a=[0:255]/256;
aa=[a' a' a'];

 figure(1),
 subplot(121),imagesc(result),colormap(gray),title('(a)result ann trained and simulated image');
 subplot(122),imhist(result,aa),title('(b)histogram')

 %save result


result1=zeros(size(result));
result1(find(result>=175))=255;
result1(find(result<175))=0;

figure(2),
imagesc(result1),colormap(gray);

%save result1

figure(3),imagesc(b),colormap(gray),title('orjinal');
figure(4),imagesc(result1),colormap(gray),title('ann simulated');

% Layer1_Weights=network.iw{1};
% Layer1_Bias=network.b{1};
% save Layer1_Bias
% Layer2_Weights=network.lw{2};
% Layer2_Bias=network.b{2};
% save Layer2_Bias
% Layer1_Weights;
% save Layer1_Weights
% Layer1_Bias;
% save Layer2_Weights
% Layer2_Weights;
% Layer2_Bias;
% Actual_Desired=[y' out'];
% Actual_Desired;
% save Actual_Desired



 testim=zeros(m,n);


for i=1:m;
```

```
   for j=1:n;
      if b(i,j)==result1(i,j);
         testim(i,j)=0;
         else testim(i,j)=255;
      end

   end
end

figure(6), imagesc(b),colormap(gray), title('target');
figure(7), imagesc(result1),colormap(gray), title('result');

figure(8), imagesc(testim),colormap(gray), title('difference between real target and result');

%calculate error
ms=length(find(b==255))%target pixel number
tp=length(find((b==255)&(result2==255)))
% h %true pixel number of the result image
w1=ms-tp %false negative pixels
ts=length(find(testim==255))

rs=length(find(result2==255))
w2=rs-tp %false positive pixels


%percent of true pixels
tt=(tp/ms)*100

%%%%%%%%%%%%%%%%%%%%
%%%%calculation of RMSE
%%%%%%%%%%%%%%%%%%%%

rmse=sqrt(tp/ms)
```