Topographic Mapping of MRI and MRS Data in Brain Tumours Assessment

Ghita BERRADA

MSc by Research in Pattern Analysis and Neural Networks

ASTON UNIVERSITY

September 2007

Abstract

About 29,000 people are diagnosed with brain tumours in Europe each year. Their prognosis and handling largely depends on the type and grade of tumour involved. The early and accurate detection and diagnosis of the tumour is very critical to the prognosis of these patients. To diagnose brain tumours, physicians can use several methods. Among these are Magnetic Resonance Imaging (MRI) and Magnetic Resonance Spectroscopy (MRS) which have received considerable attention of late as they involve almost no risks for the patients. However, these techniques generate a very large amount of data not very easily interpreted by the clinician. In this thesis, we are looking at the usefulness of applying topographic mapping on MRI and MRSI (Magnetic Resonance Spectroscopy Imaging or Multi-voxel MRS) data in assessing brain tumours and comparing the use of topographic mapping and clustering/classification on these kinds of data in brain tumour assessment.

Contents

List of Figures

List of Tables

Acknowledgements

There are so many people I would like to thank for various reasons that I am afraid to forget anyone. All I can promise is to try my best to remember everyone.

I would like to thank the Home Office for having made it so easy and effortless for me to come and stay in the UK.

I am deeply grateful to Vicky Bond for having smoothed over every administrative issue I might have encountered and for helping me find partial funding for my MSc.

I am thankful for the funding I have received from the EU BIOPATTERN Network of Excellence, under contract 508803.

I would like to thank Geert Postma from the Department of Analytical Chemistry, Institute of Molecules and Materials at Radboud University (Netherlands) for providing me with the data I have been working with for my MSc research project.

A very big 'thank-you' to my supervisor, David Lowe for his excellent guidance and everlasting good humour. I don't think I could have had a better supervisor.

I wish to thank Pierre Latouche, Dharmesh Maniyar, Diar Nasiev, Mingmanas Sivaraksa and Thomas Bermudez for all the ideas they have suggested to me and all the help on software debugging they have provided me (by which they have helped keeping me sane at times). But I want more particularly to thank Raji, Ming, Michel, Diar, Thomas, Dharmesh, Ariane, Pierre, Julien, Alex and Michael for having made my stay in Birmingham so enjoyable, so pleasant and so absolutely funny at times. Thank you for all the good times, various meals and cooking experiences as well as for all the fun we had together.

Last but not least I would like to thank my parents, my sisters Nadia and Selma and my uncle Ghali for all the love they shower on me all the time, for their endless support and for their absolute confidence in me.

A last thought for whoever put the coffee machine in the Wolfson Lab. It was particularly helpful at times to stay fully awake and operational.

Glossary

Astrocytoma

An astrocytoma is a brain tumour that grows from astrocytes ("star-shaped" cells of the brain or spinal cord. There are two types of astrocytomas:

- narrow infiltration (i.e progression) zone astrocytomas (eg. pilocytic astrocytomas,...)
- diffuse infiltration zone astrocytomas (eg. low-grade astrocytomas, anaplastic astrocytomas, glioblastomas,...) which have an intrinsic tendency to progress into more advanced grades

Low grade astrocytomas grow slowly compared to their malignant counterparts such as anaplastic astrocytoma and glioblastoma. Several years (approximately 3.5 years) can often intervene between the initial symptoms (the initial presenting symptom being seizures, often generalized, in about one half of patients) and the establishment of low-grade astrocytoma diagnosis. The clinical course of low-grade astrocytoma is in most cases either a gradual deterioration or a stepwise decline. A sudden deterioration occurs only in 15 % of the cases. The typical range of survival is more than 5 years. There is a slight male predominance for the development of this kind of astrocytomas.
Anaplastic astrocytomas are diagnosed in a mean interval of 1.5-2 years after the first symptoms. The initial presenting symptoms are usually headaches, depressed mental status and focal neurological deficits. The typical range of survival is 2-5 years. There is a slight male predominance for the development of this kind of astrocytomas.
Most cases of pilocytic astrocytoma occur between 0-20, low-grade astrocytomas between 30-40 while anaplastic astrocytomas have a median age of 40-50 and glioblastomas have a median age of 50-60.

Biopsy

A biopsy is a surgical procedure during which a sample of tissue from the suspected tumour is taken to be examined for malignancy. This procedure also helps determine the type of tumour if a tumour is detected.

Brain-blood barrier

The brain-blood barrier (BBB) is the physiological interface between the blood and the brain parenchyma (i.e the brain specialized tissue).

Brain tumour

Brain tumours are cancers that affect the brain. They can either originate from brain cells, in which case they are called primary, or be metastatic tumours (i.e resulting from the proliferation of a cancer originating from outside the brain).

The picture shows the localisation of some primary cerebral tumours in the brain.

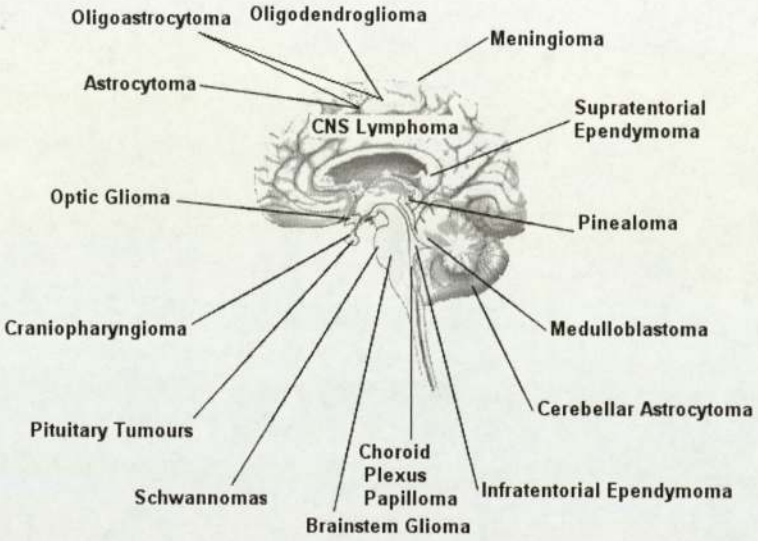


Figure 1: Some primary brain tumours and their localisation within the brain.

Choline Cho

Choline is a precursor of acetylcholine (a neurotransmitter) and a component of certain phospholipids. Choline-containing compounds are involved in the synthesis and degradation of cell membranes. In the context of detection of brain cancer, the increase of the choline-containing compounds signal intensity would mean a fast cell proliferation processus indicative of cancer.

CT Computed Tomography

CT is an imaging method in which X-ray measurements from different angles are combined into an image.

Creatine Cr

Creatine is a substance found mainly in the brain, the heart and skeletal muscle, which

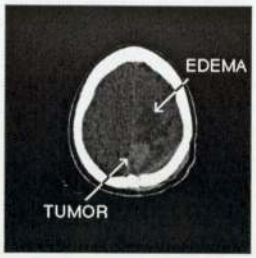


Figure 2: Contrast-enhanced CT scan of a glioblastoma multiforme (high grade tumour).

plays an important role in energy metabolism. It is phosphorylated (i.e phosphate groups are added to it) into phosphocreatine which is an energy storage compound.
^{1}H MRS detects both creatine and phosphocreatine in the so-called creatine signal.

Glioblastoma

A glioblastoma is the most malignant kind of astrocytomas. Its survival typical range is less than 1 year. There is a small male predominance in the development of this astrocytoma. It may develop from lower grade astrocytomas such as pilocytic astrocytomas (grade 1), low-grade astrocytomas (grade 2), anaplastic astrocytomas (grade 3).

Herniation

A brain herniation is the displacement of blood vessels, cerebrospinal fluid and brain tissue outside the compartments of the brain they usually occupy. It occurs when the pressure within the skull increases and displaces brain tissues. This increase in intracranial pressure is commonly the result of brain swelling following brain injury, of brain tumours, hemorrhages and strokes causing swelling within the brain. Brain herniation leads to a massive stroke due to poor supply in blood of some areas of the brain and compression of vital structures regulating the breathing and circulation. This rapidly provokes death or brain death.

The symptoms of a brain herniation are a progressive loss of consciousness, coma, irregular breathing, irregular pulse, respiratory and cardiac arrest, loss of all brainstem reflexes such as blinking, pupillary reaction to light.

Lumbar puncture

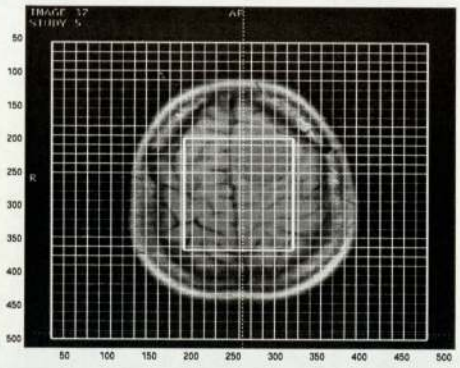
A lumbar puncture is a procedure in which a sample of spinal fluid is obtained and examined for malignancy.

Meningioma

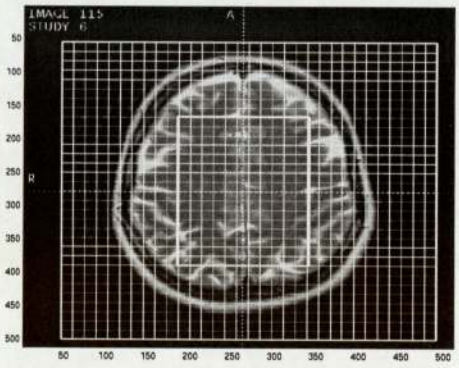
A meningioma is a brain tumour usually benign that develops in the meninges (i.e the membranes that envelop the brain and spinal cords).

MRI Magnetic Resonance Imaging

MRI is a medical imaging method that relies on the principles of atomic nuclear-spin resonance and uses strong magnetic fields and radio waves to offer sharp and precise images of the internal body structures.
Perfusion MRI is used to measure the changes in blood flow in tissue.



(a) Meningioma



(b) Healthy Brain

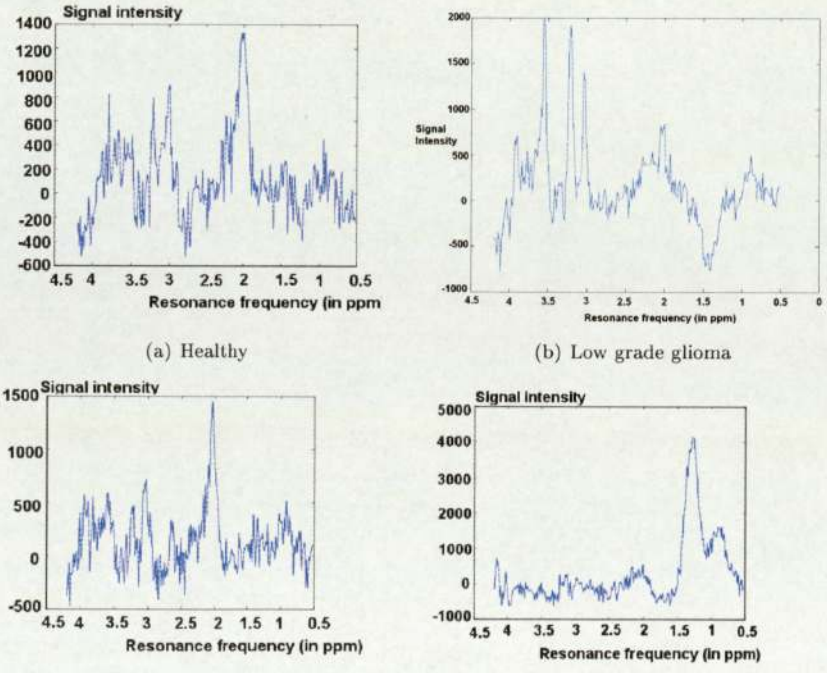
Figure 3: MRI scans.

MRS Magnetic Resonance Spectroscopy

MRS is a variation on MRI (see **MRI**) that can identify the chemical composition of a tissue.

There are two types of MRS:

- the first is localized single volume MRS, also called single-voxel MRS, in which the location to scan is previously identified using conventional MRI. The MRS signal is then taken only from the identified location. To obtain signals from other locations, the process is repeated.
- the second is magnetic resonance spectroscopic imaging (MRSI), or chemical shift imaging (CSI), in which MRS signals are acquired from a large number of voxels simultaneously.



(a) Healthy

(b) Low grade glioma

(c) Meningioma

(d) High grade glioma

Figure 4: Some MRS spectra of the brain.

N-acetylaspartate NAA

NAA is an amino-acid found in high concentrations in the neurons of the central nervous system. It is believed to be a marker of viable neurons.

Parenchyma (or parenchymal tissue)

The parenchyma, in anatomical nomenclature, consists of the tissues of an organ that perform the special function of this organ as opposed to the organ's framework called stroma.

PET Positron Emission Tomography

PET is an imaging technique that uses radioactive tracers emitting positrons (i.e particles resembling electrons except that they are charged positively) to measure cerebral activity.

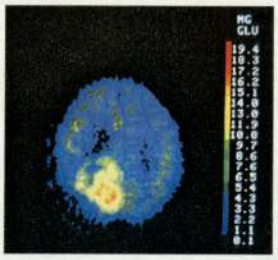


Figure 5: PET scan showing a brain tumour.
The tumour is the bright yellow and orange area at the lower left portion of the brain.

PRESS

It is a single-voxel localized MR spectroscopy technique using a spin-echo pulse sequence, that is a sequency made of 90° radiofrequency pulses to excite the magnetization and 180° pulses to refocus the spins to generate signal echoes.

Radiation Necrosis

Radiation necrosis is a long-term complication of radiotherapy or radiosurgery that takes the form of a focal structural lesion that appears on the original tumour site.

Shimming

Shimming is the process by which the main external field of an MR unit is finely adjusted to maximize its homogeneity.

STEAM Stimulated echo acquisition mode

STEAM is a single-voxel localized MR spectroscopy technique. The basic STEAM sequence is made of three selective 90°radiofrequency pulses. Each of the radiofrequency pulses is applied in the presence of a switched gradient field so as to define three orthogonal planes. The three pulses generate a stimulated echo from a single region at the intersection of these planes. As STEAM is a single shot method, it can be used to shim the signal from the region of interest so as to improve the spectral quality.

Stroma

The stroma is either:

- the connective tissue framework of an organ, gland, or other structure, as distinguished from the tissues performing the special function of the organ or part (called parenchyma)

or

- the spongy, colorless framework of a red blood cell or other cell.

Voxel

A voxel is the smallest distinguishable unit of volume, the 3D equivalent of the pixel.

Chapter 1

Introduction

1.1 Diagnosis techniques of brain tumours

A brain tumour can be diagnosed through invasive and non-invasive methods once a combination of symptoms as well as the evaluation of neurological functions suggest the possibility of a brain tumour in the patient.

Invasive methods include biopsy and lumbar puncture. Both procedures are more risky for the patient than a non-invasive diagnosis method.

Biopsy includes the risks of surgery and anesthesia (a general anesthesia is needed for this procedure), leaves a scar that can potentially cause seizures, and may be associated with brain injury as it leads to the removal of brain tissue. It does not always provide the correct answer as it is difficult to collect the tissue sample to be analysed for the existence or absence of brain tumour from the correct and exact location. Furthermore, it can only give local information, that is, at best, partial accurate information on the tumour as tumours are known to be infiltrative and heterogeneous. The use of a non-invasive imaging technique prior to a biopsy procedure can help ascertain the locations from which to collect the samples that would help give complete and accurate information on the tumour detected if any.

Lumbar puncture can cause post-lumbar puncture headaches, hematomas, etc. Lumbar puncture is usually used when the diagnosis or the type of tumour is unclear or when the tumour is suspected to have invaded the meninges and to be compressing the cranial nerves. However, this procedure is not to be performed in the case of patients having large tumours increasing intracranial pressure as the removal of cerebrospinal fluid during this procedure would cause the tumour to move thus leading to an herniation of the brain which can be fatal. Lumbar puncture and biopsy might be done after an MRI has been performed to ensure the safety of the procedures.

Non-invasive methods include PET, CT, MRI and MRS (mostly ^{1}H MRS and ^{1}H MRI) principally.

PET is not routinely used for diagnosis. It is rather used to give an indication on the grade of the tumour or to differentiate between radiation necrosis and recurring tumours. A PET scan

provides an image of the brain activity rather than its structure unlike MRI or CT. This image of the brain activity is obtained by measuring levels of injected fluorodeoxyglucose FDG (i.e glucose labelled by a radioactive tracer). It can thus assess the rate at which the tumour is using the injected glucose as compared to the rate at which this glucose is used by other parts of the brain. Typically, low grade tumours use less glucose than the healthy parts of the brain and high grade tumours use more glucose than the healthy parts of the brain. It can also be very useful in distinguishing between a recurring tumour and radiation necrosis, as cells in radiation necrosis are dead and therefore do not use any glucose while tumorous cells use glucose at a rate depending on their grade.

CT uses X-rays to provide an image of the structure of the brain but is usually less accurate in the detection of brain tumours than MRI and further includes the risks linked to the use of X-rays as well as risks of allergy due to the presence of iodine in the contrast agent used to take the images. CT has largely been replaced by MRI.

Both MRI and MRS are innocuous for the patients except patients with pacemakers or metallic parts in their bodies, as both imaging techniques rely on the use of strong magnetic fields, usually between 1.5 and 3 Teslas (about 10^5 more intense than the Earth magnetic field which is about $5 \cdot 10^{-5}$ Teslas).
A property of MRI is that stagnant blood has a signal intensity (brightness) determined by the proton density, T1, and T2 of blood, and by the contrast "weighting" chosen and appears bright in most "weightings" while circulating blood will - due to its flow velocity - in most instances induce no MR signal, and thus behave as an effective "negative" contrast medium. As most brain tumours have corrupted brain-blood barriers, this MRI property makes them appear bright, therefore making tumours, even small ones, easily detectable with MRI.
Another advantage of MRI is its spatial resolution which is about 1 mm^3 in most medical applications and up to 1 μm^3 in experimental settings as well as its good contrast resolution (higher than that of CT). These resolution properties of MRI allow a precise and sharp three-dimensional visualisation of the tumour bulk, which can prove useful in localising the area if surgery is needed.
MRS provides some information on the brain activity through the detection of some chemical substances, for example choline (which is involved in the process of synthesis and degradation of cell membranes and whose increase is considered as indicative of cell proliferation and possibly brain tumours), N-acetylaspartate (considered a marker of viable neurones), lactate (whose level increases when cells tend to use anaerobic pathways to consume glucose as is the case with tumours) and myoinositol (whose high levels indicate glial hypertrophy and proliferation that may be related to brain cancer). Both imaging techniques are described more precisely in section 1.2 on the next page.

As MRI and MRS are innocuous, both methods are being studied to determine whether their use can predict the existence of a brain tumour accurately, and even find out the type of a tumour

and therefore avoid the use of biopsy in certain cases of patients with indeterminate brain lesions and no tumour history, and differentiate recurrent tumours from radiation necrosis for patients who have undergone previous brain tumour treatments and present suspicious brain lesions.

1.2 Basic concepts of Magnetic Resonance Imaging Techniques

1.2.1 General principle of both imaging techniques

Though MRS can be performed with other nuclei than ^{1}H such as ^{13}C, ^{31}P or ^{19}F, our explanation will only focus on ^{1}H MRS also known as proton MRS as it is the most common MRS technique and provides the MRS part of our data. Both MRI (Magnetic Resonance Imaging) and MRS (Magnetic Resonance Spectroscopy) rely on the fact that protons have a magnetic moment. The patient is placed in a magnetic field ($\boldsymbol{B_0}$) as shown in figure 1.1 (in subsection 1.2.1).

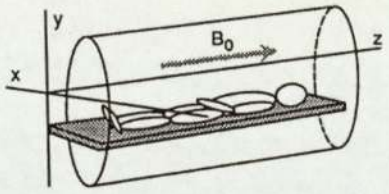


Figure 1.1: Patient during MRI (Source: *http://www.medcyclopaedia.com/library/*, n.d.).

Applying the external magnetic field $\boldsymbol{B_0}$ to the protons of the patient's tissues makes the protons' magnetic moments line up in the direction of the applied magnetic field.

In addition, the magnetic axis of each proton starts to rotate around the direction of the external magnetic field $\boldsymbol{B_0}$. This rotation, called precession, has a frequency ω_0, named the resonance frequency or Larmor frequency, proportional to the external magnetic field $\boldsymbol{B_0}$. The frequency ω_0 is given by the Larmor equation:

$$\omega_0 = g \cdot \boldsymbol{B_0}$$

where g is a constant named the gyromagnetic ratio

The gyromagnetic ratio g is specific to each type of magnetic atomic nucleus, and for the hydrogen nucleus, the ratio is equal to 42.58 MHz/Tesla. This implies that at the magnetic field strengths used in MRI (between 1.5 and 3 Teslas), the Larmor frequency of the hydrogen nucleus is in the frequency range between 63.87 MHz (at 1.5 Teslas) and 127.74 MHz (at 3 Teslas).

The magnetic moments of the precessing protons do not all point in the same direction as $\boldsymbol{B_0}$. A small majority of protons precess with their magnetic moments aiming towards the direction of $\boldsymbol{B_0}$, that is in a direction closely parallel to the external magnetic field. These hydrogen nuclei

are called "parallel protons". The remainder of the protons precess with their magnetic moments close to anti-parallel to the external magnetic field. These are the "anti-parallel protons". As a result, a net magnetic moment is created; the tissues become magnetic, and the magnetism ($\boldsymbol{M}$) is oriented exactly parallel to the external magnetic field $\boldsymbol{B_0}$. At first, the tissue magnetism has no precessional motion as, although the individual protons all precess, they are evenly distributed around the $\boldsymbol{B_0}$ direction, leaving no magnetic component in the x-y plane (see figure 1.1 on the preceding page in subsection 1.2.1 on the previous page). The size of $\boldsymbol{M}$ is determined by the surplus of parallel protons which is proportional to the external magnetic field strength, but is always very small (only in the order of 1-10 protons per 1 million protons). But $\boldsymbol{M}$ is also proportional to the number of protons per volume unit of tissue. Therefore, the enormous amount of protons present in most tissues (about 10^{22} per ml of water), accounts for the fact that the net magnetic moment $\boldsymbol{M}$ is strong enough to induce an electric current in a receiver coil located outside the patient.

If we want the net magnetic moment of the patient's tissues to induce an electric current in a receiver coil placed on the outside of the brain of the patient, we need this magnetic moment's strength running through the bore of the coil to change. This can be achieved by transmitting an electromagnetic radio frequency pulse along the patient in a direction orthogonal to the external magnetic field $\boldsymbol{B_0}$. If the frequency of the electromagnetic pulse is equal to the Larmor frequency (phenomenon called magnetic resonance), the protons' magnetic moment will change direction and so will the net magnetic moment of the patient's tissues. The strength and duration of the radio frequency pulse determine the angle of rotation of the net magnetic moment of the patient's tissues from the direction of the external magnetic field.

A receiver coil is placed on the outside of the brain with its bore oriented towards the patient, perpendicular to the external magnetic field direction. The rotation of the net magnetic moment of the patient's tissues, when an electromagnetic radio frequency pulse is transmitted, will induce an electric current in the coil, and this electric current is called the MR signal (signal represented in figure 1.2 (in subsection 1.2.1 on the previous page))

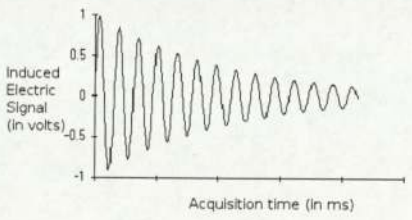


Figure 1.2: MR signal (also called free induction decay).

1.2.2 More details about MRI

Image contrast

Contrast in MR images is determined by differences in tissue magnetisms, or more precisely, by the different strengths of magnetism that rotate in the x-y plane and induce currents in the receiver coil.
Tissue magnetism is first of all determined by its proton density. Therefore, anatomic areas containing very few protons, like air, will always induce very weak MR signals and therefore always appear dark in images whereas anatomic regions with very high proton density such as water and other fluids, presumably should always appear bright in MR images. Yet depending upon to the method used for image acquisition, fluids (like CSF) may appear either bright or dark. This is because it is not only proton density that determines image contrast. Several other parameters play a role as well, and the two most important of these are named T1 and T2.

To reconstruct an image, several MR signals are needed, which implies that several radio frequency pulses must be transmitted. Between the pulse transmissions, the protons undergo two different relaxation processes called T1- and T2-relaxation.
Figure 1.2 on the preceding page (in subsection 1.2.1 on page 17) shows that the induced MR signal in the receiver coil decays quickly. This is partly due to T2-relaxation.
When the radiowaves used to induce an MR signal are no longer transmitted, the magnetic moment in the x-y plane gradually disappears because of small differences in the local magnetic field strength (partly due to magnetic molecules in the tissues) therefore inducing a decay in the MR signal. The disappearance of the component of $\boldsymbol{M}$ in the x-y plane ($\boldsymbol{Mxy}$) is due to the fact that protons exposed to slightly different magnetic field strengths will precess at slightly different Larmor frequencies, and the surplus "parallel" protons that were closely packed around $\boldsymbol{Mxy}$ immediately after the radiowave pulse will dephase (in other words spread out around the z-axis until the individual protons are evenly distributed around the z-axis), at which point $\boldsymbol{Mxy}$ is a null vector. This phenomenon is called T2 relaxation, the T2 time being defined as the time until $\boldsymbol{Mxy}$ has lost 63% of its original, maximum value.
$\boldsymbol{Mxy}$ disappears completely after 4-5 times the value of T2. The T2 value varies considerably with the physical and chemical properties of the tissues. For instance, fluid and fluid-like tissues typically have a long T2 ($\boldsymbol{Mxy}$ and the MR signal disappear slowly), while solid tissues and substances have a short T2 ($\boldsymbol{Mxy}$ and the MR signal disappear rapidly). The typical value of T2 in parenchymal tissues is about 50 ms.
T1 relaxation is a slower process than T2 relaxation, and is the phenomenon in which the protons' magnetic moments regain the direction they had before a radio pulse was applied, that is a direction parallel to $\boldsymbol{B_0}$. During this process, the net magnetic moment along the z-axis, $\boldsymbol{Mz}$, will increase from zero until it reaches its maximum value, determined by the proton density in the tissue. The T1 time is therefore defined as the time until $\boldsymbol{Mz}$ has regained 63 % of its original, maximum value. The shorter the T1, the faster the restoration of $\boldsymbol{Mz}$. Generally the

original value of Mz is regained after 4-5 times the T1 value. The T1 value is largely determined by molecular size and mobility and varies largely between different types of tissues. Generally, T1 is shortest in tissues having molecules of medium size and mobility such as adipose tissue while smaller, more mobile molecules (as in fluids) as well as larger, less mobile molecules (as in solids) have longer TI values. A typical T1 value in parenchymal tissue is approximately 500 ms. Knowing that, the operator of a MR unit may decide whether image contrast should be determined mainly by proton density, T1 or T2 by adjusting the time period between the transmission of radio frequency pulses.
The longer the time interval (up to a certain point) between two pulses, the larger the value of Mz regained, the larger the Mz to be rotated into the xy-plane by the next pulse, and therefore the stronger the MR signal induced. If the next pulse is transmitted before the T1 relaxation in the tissues is completed, the value of Mz in the tissues will depend upon their T1 values. Assuming the tissues have fairly similar proton densities, tissues having the shortest T1 will have regained the largest Mz, and will consequently induce the strongest MR signals after the following pulse and therefore appear bright in the final image. Similarly, tissues with the longest T1 will induce the weakest signals. The images obtained through this process are called T1-weighted images. T1-weighted images are acquired with relatively short repetition times (TR) (that is intervals between two successive pulses) typically about 500 ms.

Alternative image contrasts can be obtained by increasing TRs, be it proton density weighted images or T2-weighted images. In proton density (PD) weighted images, the tissues with the highest proton density induce the strongest MR signals and appear bright while in T2-weighted images, the brightest tissues are those having the longest T2. Long TRs are needed to obtain both types of contrasts so as to suppress the effect of differences in T1 on image contrast. To ensure that all tissues have completely regained their maximum Mz before transmission of the next pulse, TR values more than 5 times the longest T1 of the tissues are used.
The maximum magnetism along the z-axis being determined by the proton density, the relative strengths of the MR signals derived from the various tissues immediately after the pulses will reflect the relative proton densities of the tissues. The contrast obtained thus is called PD-weighted (proton density weighted).
T2-weighted contrast is achieved by introducing a time interval (called echo time, TE) between the pulse and the signal measurement. During this time interval, T2-relaxation induces a reduction of the size of Mxy; slow reduction in tissues having a long T2, quicker reduction in tissues having a short T2. The amplitudes of the induced MR signals recorded at the end of the echo time will therefore reflect the differences in T2 in the tissues.
Because of all these contrast techniques, image contrast in MR imaging can be made much more variable than image contrast in alternative modalities such as computed tomography. Image contrast being determined by operator-dependent parameters such as repetition time and echo time, and by tissue-dependent parameters such as proton density, T1, and T2, a basic knowledge

of these parameters and their modulation can provide a means of detecting the kind of tissues a clinician is interested in and for example allow for better contrast for soft tissues than can be obtained with CT.

Slice selection

Induced MR signals are produced by the transmission of radio frequency pulses only when the pulse frequency is exactly equal to the proton Larmor frequency. This makes it possible to collect MR signals from a predetermined thin slice of tissue. The first step towards slice selection is to create a magnetic field gradient through the anatomical region to be imaged. Special coils (called gradient coils) create small additional magnetic fields, which results in the strength of the $\boldsymbol{B_0}$ magnetic field increasing linearly in one direction. The Larmor frequency of the protons being proportional to the magnetic field strength (see the Larmor equation 1.1 on page 17), the Larmor frequency will therefore increase linearly in the gradient direction. By transmitting radio frequency pulses having a predetermined narrow frequency range, MR signals will be recorded from only the thin slice of tissue that has a Larmor frequency range corresponding to the pulse frequency range. The orientation of the magnetic field gradients and consequently also the slice directions, can be freely selected.

Image reconstruction

The tissue slice to be imaged may be considered as consisting of several equally large volume elements, called voxels. After each radio frequency pulse, every voxel has its own tissue magnetism ($\boldsymbol{Mxy}$) which induces a signal in the receiver coil. The amplitude of the individual voxel signal is determined by the size of the voxel tissue magnetism, which depends on voxel dependent factors such as proton density, T1, and T2, as well as the choice of repetition time and echo time. Each voxel corresponds to a picture element, pixel, in the final two-dimensional image. The brightness (shade of grey) of the pixel is determined by the signal amplitude induced by the magnetism in the corresponding voxel.

As the MR signal recorded from a slice of tissue is a composite signal induced by all the individual voxel magnetisms simultaneously, there is a need to differentiate between the various voxel signals to assign the correct brightness to each pixel. The differentiation between voxels is done using the frequency and phase of each voxel signal, which is determined by the frequency and phase of the rotating voxel magnetism $\boldsymbol{Mxy}$. The encoding of the voxels is obtained by applying magnetic field gradients in the y- and x-direction, respectively (for axial slices). The gradients affect the rotation of the voxel magnetisms in such a way that the resulting voxel signals have a phase that depends on the voxel location along the y-axis, and a frequency that depends on the voxel location along the x-axis. This ensures that each voxel is given a unique code of phase and frequency and can be differentiated from the other voxels of the slice.

1.2.3 More details about MRS

The underlying principle of MRS is that atomic nuclei are surrounded by a cloud of electrons which very slightly shields the nucleus from any external magnetic field. The structure of the electron cloud being specific to an individual molecule or compound, the magnitude of this screening effect is a characteristic of the chemical environment of individual nuclei. As the resonance frequency of protons is proportional to the magnetic field that it experiences (see the Larmor equation 1.1 on page 17), the resonance frequency will be determined not only by the external applied field, but also by the small field shift generated by the electron cloud. This shift in frequency is called the chemical shift. Chemical shift is a very small effect, usually expressed in "parts per million" of the main frequency. Therefore in order to resolve different chemical species, it is necessary to achieve very high levels of homogeneity of the main magnetic field $\boldsymbol{B_0}$. The main disadvantage of proton MRS is that signals from metabolites are several orders of magnitude smaller than the signal from water and, if present, fat. This, coupled with the fact that the proton chemical shift dispersion is rather small (approximately 10 ppm) means that very good shimming and exceedingly efficient water suppression are required. Strong proton signals come from N-acetylaspartate (NAA), creatine plus creatine phosphate (Cr), choline-containing compounds (Cho) and myo-inositol. In some pathologies (for example high grade brain tumours), a lactate signal can be observed. The amount of information in proton spectra highly depends on the experimental conditions. In particular, if a spin- or a stimulated-echo sequence is used (STEAM and PRESS) then the echo time determines the nature of the spectrum: at long echo times (for example 270 ms) one observes simple spectra with only a few peaks whereas at short echo times, the spectrum has an improved signal to noise ratio (SNR), but is much more complicated as overlapping broad signals from many moderate to large molecules (for example proteins) are observed, which distorts the apparent baseline and makes it difficult to determine the concentration of each metabolite for which a resonance peak appears on the spectrum.

1.3 Use of MRI and MRS for brain tumours diagnosis

The usefulness of MRS and MRI has been so far assessed in three main tasks: brain tumour detection (that is differentiating between healthy and tumorous tissues), grading of brain tumours and differentiation between recurring brain tumours and radiation necrosis. We will not review the studies done on the last task as we are primarily interested in the first two ones.

1.3.1 Brain tumour detection

Conventional MRI does not assess the presence and extent of brain tumours precisely because of practical limitations (Kurhanewicz, Vigneron, and Nelson, 2000). Visualizing the presence of a tumour through MRI is based on the assumption that most brain tumours are characterized by a compromised brain-blood barrier and consequently enhance on the MRI images after injection of a contrast substance (for example gadolinium). On those images, the volume enhanced by the contrast agent is identified as the tumour bulk. But usually the contrast-enhanced volume does not give an accurate reflection of the true extent of the tumour because some tumour

tissue does not enhance with contrast agent injection and because contrast-enhancing necrosis can appear. A review of 100 patients (by Nelson SJ, Vigneron DB, Dillon WP(1999) quoted by John Kurhanewicz et al. (Kurhanewicz et al., 2000)) showed a significant number of patients (66%) presented abnormal metabolism in 3D MRSI that could indicate a tumour area outside the contrast-enhancing area obtained through contrast-enhanced MRI. Therefore, some studies were made to determine whether proton MRS could provide an accurate differentiation of tumours and normal tissues.

The studies showed that normal brain tissues presented dramatic spectral differences compared with tumorous tissues: the intensity of NAA signals was reduced in tumorous tissues compared with normal ones and the choline peak was higher than normal in tumours. Creatine was also reduced and lactate reached high levels in tumours but these two results were not significant as they were variable. The loss in NAA was conjectured to be due to neuronal cell density decrease in tumour bulks whereas choline was linked to the proliferative nature of tumorous cells as it is a compound involved in the generation of the cell membranes.

High choline intensities, decreased levels of NAA and creatine compared with normal tissues were again shown to be a characteristic of tumours in a multisite, 15-institution trial of single-voxel MRS conducted on 86 patients (1996) (quoted by John Kurhanewicz et al. (Kurhanewicz et al., 2000)).

Other studies investigated the significance of metabolic ratios in order to provide a quantitative index for tumour discrimination. Asavaphatiboon et al. (Asavaphatiboon, Sinlapawongsaa, Laothamatasa, Dhanachaia, Theerapancharoenb and Putthicharoenratc, 2004) compared the choline/creatine (Cho/Cr), N-acetylaspartate/creatine (NAA/Cr) ratios (obtained through MRS) between benign and malignant infiltrative brain tumours in a sample of 9 patients with histological diagnosis of benign and malignant tumours. They showed that the presence of a tumour was characterized by an increased Cho/Cr ratio and a decreased NAA/Cr ratio. Therefore MRS could indicate the presence of a tumour, though knowing whether it was benign or malignant remained challenging. However the sample was small, therefore the generalisation of the study results remains questionable.

Another kind of studies (Devos, Simonetti, van der Graaf, Lukas, Suykens, Vanhamme, Buydens, Heerschap and van Huffel, 2005, Lukas, 2003, Simonetti, 2004, Devos, 2005) focused on the use of classifiers to differentiate healthy and tumorous tissue. For these studies, several kinds of data were compared: MRSI full spectra, MRSI peak integration values, MRI image intensities and the combination of MRS peak integration values and MRI image intensities. The classifiers tested were Linear Fisher Discriminant (with PCA for dimensionality reduction), Least Squares vector machines with linear kernel and RBF kernel, Gaussian Mixture Models. The area under the Receiver Operating Characteristic (ROC) curve (AUC) (see appendix B on page 100) was used as a global performance measure on the test data. The use of peak integration values combined with MRI image intensities or not was found to give high performances for all techniques

(AUC higher than 0.9). These studies also concluded that all the classification techniques performed very well especially when based on the combination of peak integration values and MRI data. All this seems to suggest that MRS could potentially enable an accurate diagnosis of the presence of a brain tumour.

1.3.2 Detection of a particular kind of brain tumour

Another question that has arisen in the use of MRS and MRI in assessing brain tumours is their capacity to differentiate different kinds of tumours (once again in order to avoid invasive methods as much as possible and replace them by non-invasive innocuous methods).

The multisite, 15-institution trial of single-voxel MRS (1996) quoted from (Kurhanewicz et al., 2000) in subsection 1.3.1 on page 22 also tried to differentiate several kinds of tumours using MRS spectras. On average, astrocytomas and anaplastic astrocytomas had the highest choline peaks whereas glioblastomas had the lowest creatine intensities. But the study also found that such metabolic characteristics could vary widely within each subtype of glial tumours and the overlap resulting from it made the distinction between low and high grade tumours very inaccurate.

Yet in a study made in 1996, Preul et al. (Preul, Caramanos, Collins, Villemure, Leblanc, Olivier, Pokrupa and Arnold, 1996) showed that a pattern-recognition analysis of the biochemical information obtained through proton MRS could allow an accurate noninvasive diagnosis of the most prevalent brain tumour types in adults.

In some studies based on the use of classifiers and described above in subsection 1.3.1 on page 22 (Simonetti, 2004, Devos, 2005, Devos et al., 2005) the following classes were considered :

- grade 2 gliomas
- grade 3 gliomas
- grade 4 gliomas
- meningiomas

With these classes the following classifiers are formed:

- grade 2 gliomas versus grade 3 gliomas
- gliomas (grade 2, grade 3 and grade 4) versus meningiomas
- low grade gliomas (grade 2, grade 3) versus high grade gliomas (grade 4)

Lukas' thesis (Lukas, 2003) considers the following classes as well: metastases, astrocytomas grade 2, meningiomas, glioblastomas, with which he constructs the following classifiers: meningiomas versus metastases, meningiomas versus astrocytomas grade 2, metastases versus astrocytomas grade 2, glioblastomas versus meningiomas, glioblastomas versus astrocytomas grade 2 and glioblastomas versus metastases.

The results of linear discriminant analysis (LDA), Least-Squares vector machines (LS-SVM) with linear kernel and with RBF kernel are compared in (Devos et al., 2005). The use of peak integration values combined with MRI image intensities or not was found to give high performances for all techniques: the AUC was superior to 0.91, 0.91, 0.99 for high versus low grade tumours, high versus low grade gliomas, gliomas versus meningiomas respectively when using peak integration values along with MRI data. Another conclusion was that in all cases, all the classification techniques performed very well especially when based on the combination of peak integration values and MRI data. Lukas (Lukas, 2003) concluded using the same techniques as the ones used above that the mean AUC for all classification techniques was greater than 0.90 for meningiomas versus metastases, meningiomas versus astrocytomas grade 2, metastases versus astrocytomas grade 2, glioblastomas versus meningiomas and glioblastomas versus astrocytomas grade 2. Only glioblastomas versus metastases had an AUC less than or equal to 0.6 depending on the data.

Therefore the ability of magnetic resonance imaging methods in grading brain tumours seems promising but remains however largely to be explored.

1.4 Overview of the data used

The data that has been used in this thesis has been provided by Radboud University (Netherlands) and contains the data collected from 25 patients with brain tumours (3 patients with meningioma and 22 patients with glial tumours) and 4 healthy volunteers. The glial tumours are divided into three classes, according to the rules of the World Health Organization: grade II (10 cases), grade III (4 cases) and grade IV (7 cases) tumours. The grade and type of tumour of one of the patients with glial tumours are not clear, therefore only 21 of the 22 patients with glial tumours have a grading associated to the tumour they have.

Our data is furthermore divided into a training set (described more precisely in section 1.4.2 on page 27) and patient specific data (see section 1.4.3 on page 28). The patient specific data consists in all the voxels acquired through MRI and MRSI for each of the 25 patients and 4 volunteers. The voxels of the training set represent a small sub-sample of the patient specific data, that have been selected to form the training set because they are histologically representative of certain types of tissue and tumours.

Furthermore, healthy brain tissue, and cerebrospinal fluid (CSF) classes have been constructed (in the training set). The data for healthy tissue has been selected from 4 healthy volunteers and from the contra-lateral brain region of 4 patients. Data points for the CSF have been selected from 8 patients. Only CSF voxels that were not in close contact with the tumour region have been selected. The data is derived from MRI and proton MRSI (that is multi-voxel MRS) which are relatively recent tools for cancer tumour analysis. In the following subsections, we describe the acquisition and pre-processing of the data (patient-specific and training set) then detail the composition of the training set and the patient-specific data.

1.4.1 Acquisition and pre-processing of the data

Acquisition of the data

The MR data from each patient were acquired in compliance with an acquisition protocol set up by the INTERPRET consortium. This protocol consists in:

- performing all measurements on a 1.5 T Siemens Vision whole body system, using a CP-head coil;
- acquiring conventional T1-weighted, T2-weighted and proton density weighted images then acquiring a T1-weighted image after intravenous administration of Gadolinium-DTPA (Gd image) (bolus injection of 15 ml 0.5 M Gd-DTPA), and finally performing water suppressed and unsuppressed proton MR Spectroscopic Imaging (^{1}H-MRSI).

The following parameters were used for the collection of transversal MR images: 23 slices of 5-mm thickness, inter-slice distance 1.5 mm, FOV (field of view) 230 mm, matrix 256x256. A standard spin echo sequence was used for the T1-weighted images (TE/TR=15/644 ms) while the PD (proton density weighted) image (TE/TR=16/3100 ms) and the T2-weighted images (TE/TR=98/3100ms) were simultaneously acquired using a turbo spin echo sequence. Only the four images (T1, T2, PD and GD) acquired from the same location as the MRSI were retained for the training dataset, and images just below or above were discarded from the training set, as they were not totally within the MRSI slice. The MRSI spectra were acquired using a 2D STEAM 1H-MRSI sequence with short echo time. The STEAM box was positioned in a transversal plane through the brain showing the largest Gadolinium-DTPA enhanced tumour area in the Gd image and was placed entirely in the brain parenchyma to avoid leakage of disturbing signals from fatty tissue surrounding the skull. The MRSI acquisition parameters were: 16 x 16 x 1024 samples, TR/TE/TM = 2000 or 2500 ms/ 20 ms / 30 ms, slice thickness = 12.5 or 15 mm, FOV = 200 mm, spectral width = 1000 Hz and NS = 2 (1 for the MRSI without water suppression).

Pre-processing steps

The MRI and MRSI data were processed in the same unique way for each case and automated where possible.

- **First step**
  For the suppressed (metabolite) MRSI data, K-space data was filtered by a Hanning filter of 50% using the LUISE software package (Siemens, Erlangen, Germany), followed by zero filling to 32 X 32 and 2D Fourier transformation to obtain time domain signals for each voxel.
- **Second step**
  Each voxel within the STEAM box was corrected for eddy current effects in the spectra using a method which prevents the occasional occurrence of eddy current correction induced artifacts described in Simonetti, 2004, chapter 3.

- **Third step**
  HLSVD filtering was performed to remove the residual water signal between 4.3 and 5.5 ppm.

- **Fourth step**
  Short echo time ^{1}H MRSI signals are characterized by the presence of a partially unknown broad baseline underlying the sharper resonances of the metabolites of interest, that hinders the assessment of the intensity of low weight metabolites. Therefore, a baseline correction with a filter width of 5ms was performed as follows: the broad resonances peaks are assumed to be mostly contained in the first part of the spectrum, therefore each spectrum is multiplied by a fast decaying exponential function (the result of this multiplication is supposed to contain the broad resonance peak), then the result of the multiplication is subtracted from the original spectrum.

- **Fifth step**
  The zero order phase is in principle removed by the eddy current correction, the only remaining phase correction to perform is a first order phase correction. The first order phase of the mean spectrum as calculated from all spectra within the STEAM box of each patients MRSI dataset was manually optimized. The correction value thus obtained was then applied to each separate time domain signal of that dataset.

The data obtained from the original data through all these pre-processing steps is a set of preprocessed free induction decays (FIDs) for each voxel within the STEAM box of the patient. Finally, each time domain signal was Fourier transformed to obtain a spectrum from which only the region between 0.5 to 4.2 ppm was retained (242 points).
The MR images were aligned with respect to each other, by successively shifting the T1- weighted and Gd image with respect to the PD image until a maximum of spatial correlation was reached. The MRSI spectroscopic grid was assumed to be aligned with the PD image, since they were acquired consecutively. After alignment of the MR images, the image pixels which did not fit within the boundary of the STEAM box were discarded from the training set. Then, the resolution of the remaining part of the MR images was lowered to the resolution of the MRSI grid (to help combine spectroscopic information from MRSI and information from MR images). This resolution lowering was performed by averaging the image pixels which were covered by each spectroscopic voxel and then range-scaling the values from each image in agreement to the range of the spectral data. After preprocessing, each voxel within the grid is then represented by a spectrum of 242 variables (the region between 0.5 to 4.2 ppm) or a set of quantified peaks (10 variables), and 1 variable from each MRI image.

1.4.2 The training set

The training set we use in this thesis is composed of 669 voxels sampled from across all patients, and consists of both MRI data and MRS data.

Table 1.1: Metabolites of interest and their resonance frequencies

Metabolites	Resonance frequency
first creatine/phosphocreatine peak	3.95 ppm
first glutamate peak	3.65 ppm
myoInositol	3.56 ppm
glutamate and glutamine	3.44 ppm
choline	3.20 ppm
second creatine/phosphocreatine peak	3.02 ppm
second glutamate peak	2.20 ppm
N-acetylaspartate	2.02 ppm
lactate	1.33 ppm
fatty acids	0.90 ppm

The MRI data can be split into four different channels per voxel which correspond to the MRI image intensities of four different kinds of MRI scans (T1-weighted scan, T2-weighted scan, proton-density-weighted scan, and a Gadolinium-enhanced T1-weighted scan).

The MRS data has two main feature representations:

- full spectra in the frequency range 0.5 ppm to 4.2 ppm, and
- integrated peak intensity values.

The full spectra are represented by 242 variables corresponding to the spectral points. The integrated peak values are 10 variables that are obtained by integrating the whole spectra within a frequency range of 0.26 ppm centered around the resonance value of 10 specific metabolites of interest and roughly measure the concentration of each of the metabolites of interest. The resonance values of interest are shown in the table 1.1.

1.4.3 The patient-specific data

The patient-specific data is composed of the collected intensities and spectra of MRI and MRS for a specific patient. An individuals' MRI data contain up to 65556 data points per scan and the MRS data contain between 88 and 225 data points per patient depending upon which voxels have been selected.

1.4.4 Cases considered for discrimination

We have considered several possible discriminations of the data:

- healthy tissue (318 points) versus tumorous tissue (351 data points)
- gliomas (303 points) versus meningiomas (48 points)
- low grade gliomas (grade 2 and grade 3 gliomas) (233 data points) versus high grade gliomas (glioblastomas multiforme) (70 data points)

- low grade tumours (low grade gliomas and meningiomas)(281 data points) versus high grade tumours (glioblastomas multiforme) (70 data points)
- grade 2 gliomas (136 data points) versus grade 3 gliomas (57 data points)
- astrocytoma diffuse grade 2 (60 data points) versus other grade 2 gliomas (86 data points)
- oligoastrocytoma grade 2 (45 data points) versus other grade 2 gliomas (131 data points)
- oligoastrocytoma grade 3 (28 data points) versus other grade 3 gliomas (29 data points)
- oligodendroglioma grade 3 (25 data points) versus other grade 3 gliomas (32 data points)

1.5 Plan of the thesis

In chapter 2 on the following page, we discuss why using topographic mapping can provide complementary information to the information provided with a classification approach. Then in chapter 3 on page 32 we give an account of the visualisation experiments we have done. Chapter 4 on page 60 explains the classification experiments done on the data. Chapter 5 on page 79 compares the results of the visualisation experiments with those of the classification experiments. Finally, chapter 6 on page 95 gives a conclusion on the work done and indicates further research directions.

Chapter 2

Topographic mapping versus classification

2.1 Defining the topographic mapping approach and the classification approach

Topographic mappings are a class of data-processing procedures that aim at representing high-dimensional data in a lower dimensional space while preserving the "geometric structure" of the data, that is the distance relationships between the data points being visualised: two data points lying close together (respectively far from each other) in the high dimensional space will also lie close together (respectively far from each other) in the visualisation space. Therefore, topographic mappings rely on pairwise relationships between data points to represent the data onto a lower dimensional space than the original data space. Consequently, when visualising a plot obtained through a topographic mapping method, we can infer many important relationships between data points such as the existence of clusters that is the existence of sets of points closely grouped together in the original data space which should therefore appear closely grouped together as well in the visualisation space (Tipping, 1996).

In contrast, in a classification problem, we consider a set of P data points and a set of L labels and aim at assigning to each of the P data points one of L labels. What we are really interested in, in this case, is to find out which properties characterize each of the classes -labelled with one of the L labels- and then assign each of the P data points to one of the defined classes because the properties they exhibit are those of the class they are assigned to.

2.2 Why exploiting pairwise relationships might prove useful in the brain tumours problem

Two voxels belonging to the same class (tumorous, healthy, low grade tumour, high grade tumour,...) should have chemical properties (described by MRS data) and/or anatomical properties (described by MRI data) that are similar. This similarity can be measured in terms of distances. Therefore measuring distances in the visualisation space can allow us to distinguish clusters of data points.

If we choose a reference point for the distances measured, we can by looking at the distances know how far a certain data point is from the reference. In a two-class problem, if we set the reference point to be a representative of one of the two classes, for example the median of the class (instead of the mean of the class to avoid the problems related to the non-robustness of the mean to outliers), we assess how far a given point is from one class. If we take the example of the two-class problem healthy versus tumorous and set the reference point to be the median of the healthy class (while considering the MRS data), the distance measured between a given point and the reference point gives us an indication on the deviation of this given point from "normality", that is from being healthy, and as such can give us an indication on how its chemical properties compare to the chemical properties of the healthy class and can thus help us spot anomalous points that may be classified as healthy but could potentially evolve into tumorous voxels.

2.3 NeuroScale versus other visualisation techniques

One of the aims of our study is to map back non-labelled patient-specific MRS data onto the patient's MRI scan so as to exploit the complementarity of MRS and MRI data to assess the existence and extent of a brain tumour. To achieve this, we need the topographic mapping technique used to have two main properties:

- the preservation of the global structure of the data so that measuring distances between data points reflects the "real" geometric relationship between these data points
- the ability to project new data on pre-existing maps

The topographic nature of SOM maps relies on the mechanism of the local neighbouring function used to generate the SOM maps and there is no explicit retention of the global structure of the data. In fact, because of the fixed topology of the neurons lattice, a topographic distortion of the data mapped is necessarily introduced, and in all but trivial situations "global topology distortions are [...] inevitable" as reported from Li, Gasteiger and Zupan (1993) in (Tipping, 1996). As there is no transformation defined from the input space to the feature space in Sammon maps and MDS schemes, the projection of new data points in pre-existing maps cannot be done without expensively re-generating the entire map from the augmented dataset. In contrast, NeuroScale provides a transformation of the data (using RBF networks) that allows for the projection of new points in previously generated maps and guarantees an optimal preservation of the global structure of the data as it relies on the minimisation of a STRESS function similar to that used in Sammon mappings (see appendix A.1 on page 98).

Chapter 3

Visualisation for the assessment of brain tumours

3.1 Methods used

The visualization methods we have been testing are PCA, NeuroScale and GTM.

PCA

Depending on the cases, only 2 or 3 principal components were kept after analysis of the eigenvalues as the eigenvalues showed a sharp decrease after the second or third value.

NeuroScale

The output dimension was taken to be 2. The data used to perform the visualisation is either full spectra data, peak integration data, MRI image intensities data or the combination of peak integration values and MRI image intensities, with and without normalization. Unless otherwise specified, normalizing the data means that for each variable, we subtract the mean of the variable from each component of the variable and then divide each component of the variable by the standard deviation of the variable. We use as many RBF centres as mapped data points to guarantee the smoothness of the topographic transformation.

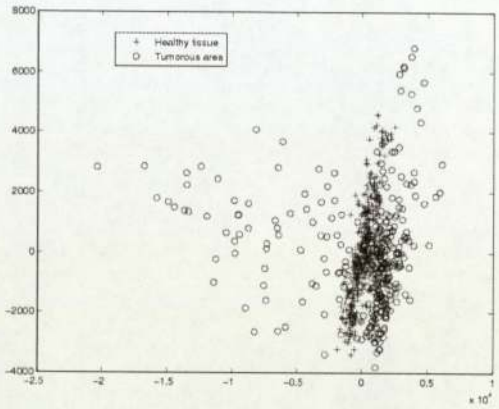
GTM

As with NeuroScale, the output dimension was taken to be 2. All the data which has been used to perform the visualisation has been normalized before applying GTM. GTM has been applied on peak integration data, MRI image intensities data and the combination of peak integration values and MRI image intensities. It couldn't be applied on full spectra data because the number of data available was too small in comparison with the number of variables of the data (242) for a GTM model to fit correctly to the data (after a cycle or two the GTM procedure stopped because the error of the model did not decrease).

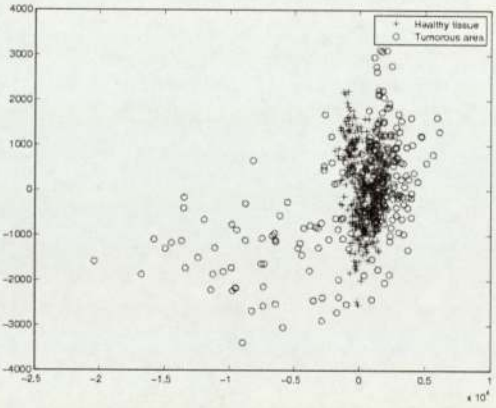
3.2 Results from preliminary visualisations

3.2.1 PCA

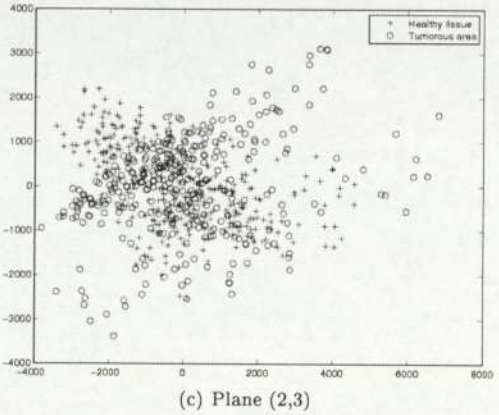
PCA is used as a benchmark vizualisation technique. Overall, in all the classifications considered, the classes showed very strong overlap and do not seem to be linearly separable.



(a) Plane (1,2)

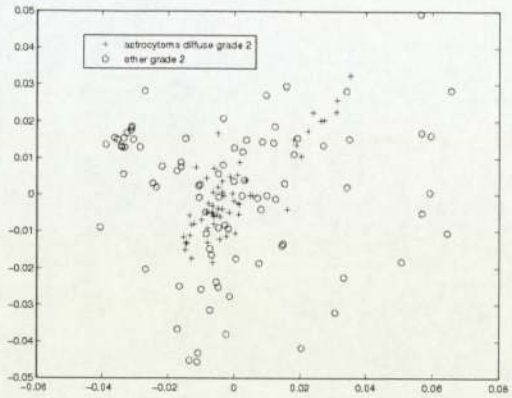


(b) Plane (1,3)

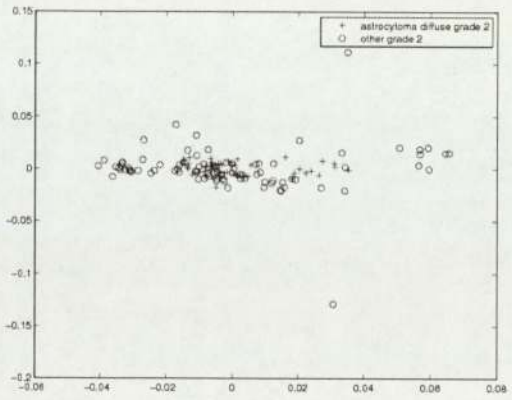


(c) Plane (2,3)

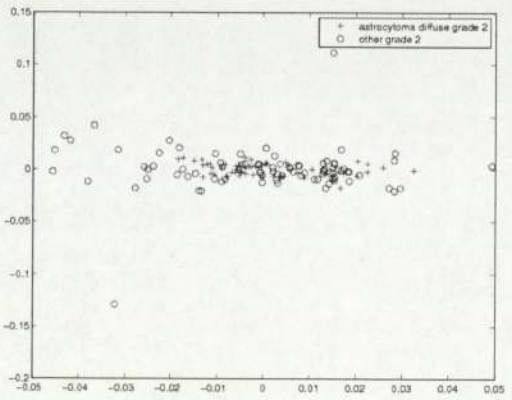
Figure 3.1: Projection of the healthy and tumorous tissue classes on the first three eigenvectors (full spectra data).



(a) Plane (1,2)

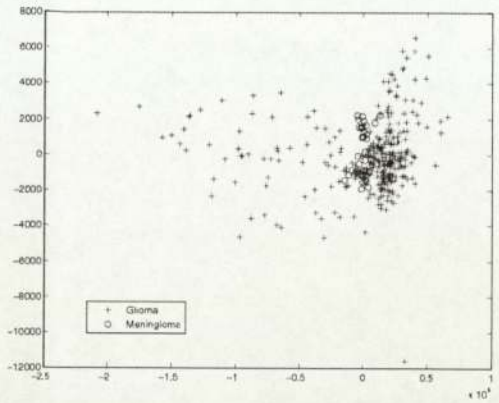


(b) Plane (1,3)

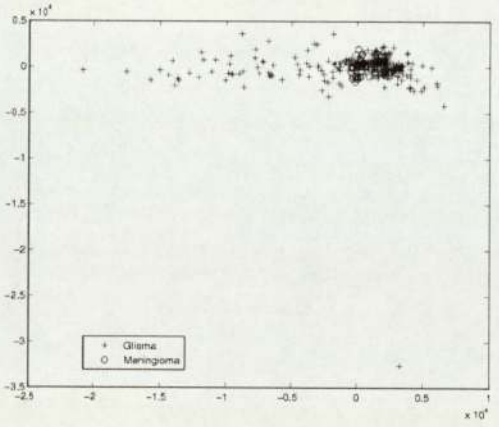


(c) Plane (2,3)

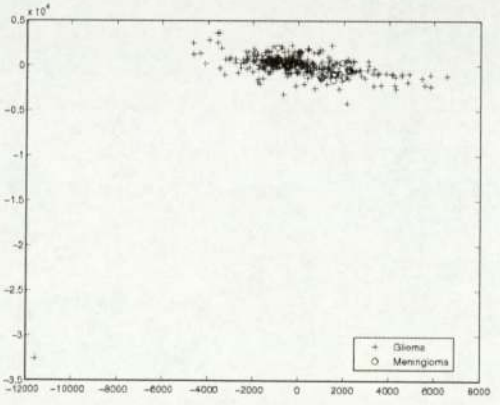
Figure 3.2: Projection of the astrocytoma diffuse grade 2 and other grade 2 classes on the first three eigenvectors (full spectra data).



(a) Plane (1,2)



(b) Plane (1,3)

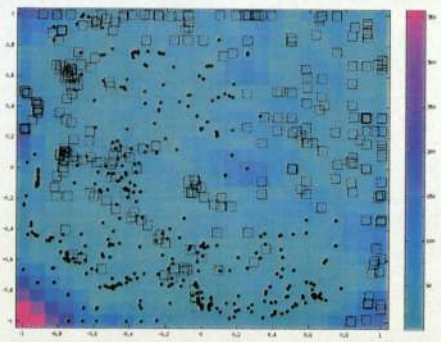


(c) Plane (2,3)

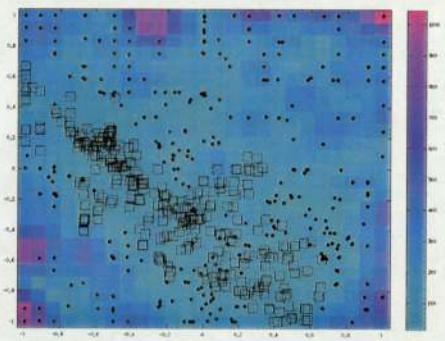
Figure 3.3: Projection of the gliomas and meningiomas classes on the first three eigenvectors (full spectra data).

3.2.2 GTM

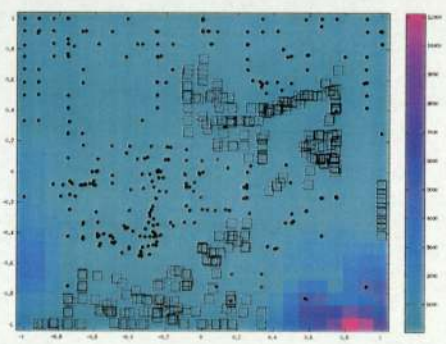
None of the cases exhibits linear separability. Overall, for almost all types of data and almost all the cases, the classes seem barely separable. Using MRI data or the combination of MRI and MRSI data seems to improve the separability in the low grade tumour versus high grade tumour case as well as the low grade glioma versus high grade glioma case (these classes could be easily separated using a non-linear classifier). Using MRI data or MRSI data improves the separability in the oligodendroglioma grade 3 versus other grade 3 case but this separability should be checked further as this could only the effect of having too few samples for this case.



(a) With MRI data

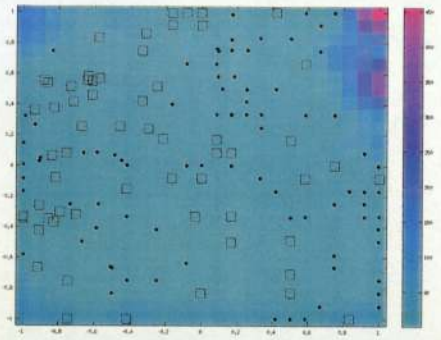


(b) With peak integration values

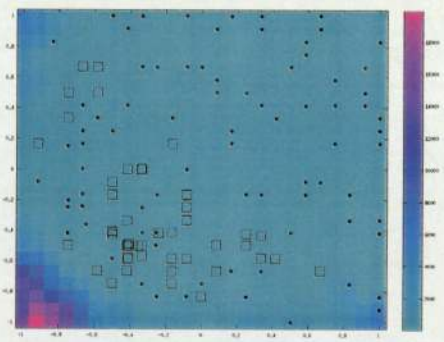


(c) With peak integration values and MRI image intensities data

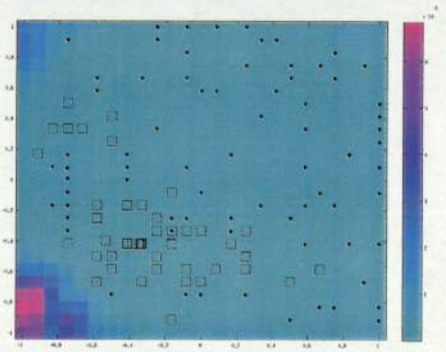
Figure 3.4: GTM with magnification factors for the healthy versus tumorous case. The dots represent the healthy class and the black squares the tumorous class.



(a) With MRI data

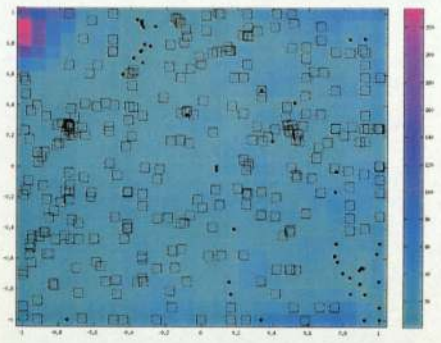


(b) With peak integration values

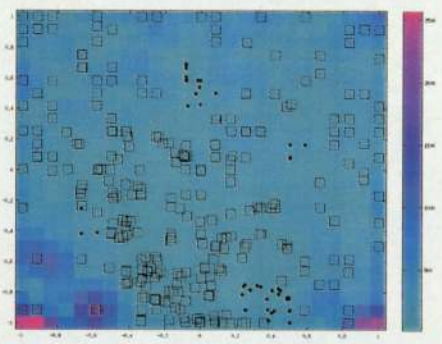


(c) With peak integration values and MRI image intensities data

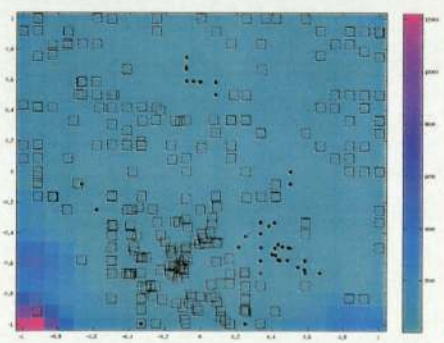
Figure 3.5: GTM with magnification factors for the astrocytoma diffuse grade 2 versus other grade 2 case. The dots represent the astrocytoma diffuse grade 2 class and the black squares the other grade 2 class.



(a) With MRI data

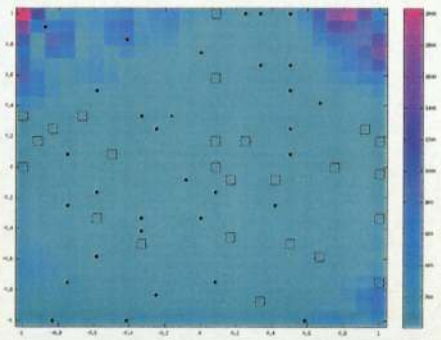


(b) With peak integration values

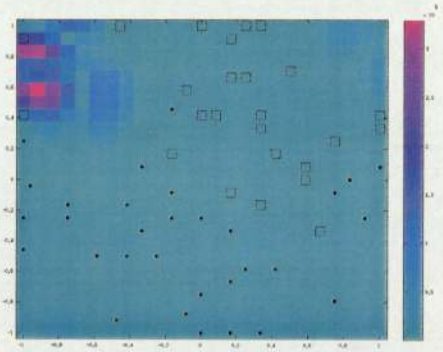


(c) With peak integration values and MRI image intensities data

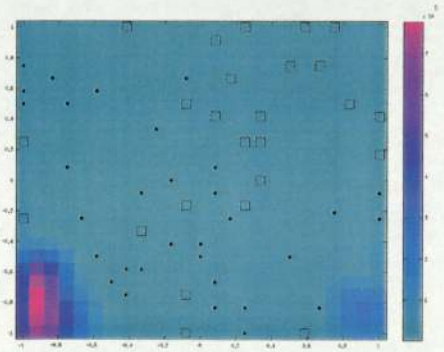
Figure 3.6: GTM with magnification factors for the glioma versus meningioma case. The dots represent the glioma class and the black squares the meningioma class.



(a) With MRI data

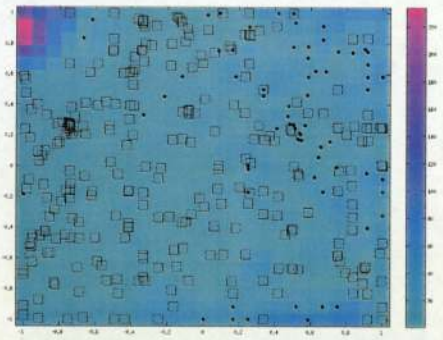


(b) With peak integration values

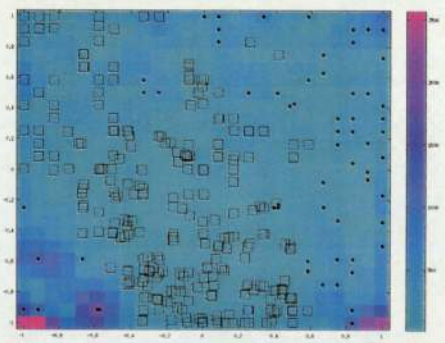


(c) With peak integration values and MRI image intensities data

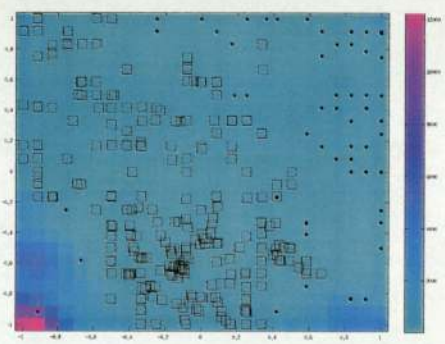
Figure 3.7: GTM with magnification factors for the oligodendroglioma grade 3 versus other grade 3 case. The dots represent the oligodendroglioma grade 3 class and the black squares the other grade 3 class.



(a) With MRI data



(b) With peak integration values



(c) With peak integration values and MRI image intensities data

Figure 3.8: GTM with magnification factors for the low grade tumours versus high grade tumours case. The dots represent the low grade tumours class and the black squares the high grade tumours class.

3.2.3 NeuroScale

All the NeuroScale models described in this thesis rely on thin plate spline radial basis functions as the STRESS function values obtained with these functions are lower than those obtained with $r^4 \log r$ functions (a hundred times lower in general). Furthermore, using Gaussian basis functions and $r^4 \log r$ functions leads to suboptimal representations of the data as the NeuroScale training stops at the first cycle for the Gaussian basis functions and after the second cycle for the $r^4 \log r$ functions as the STRESS function values do not decrease. This means that the configurations obtained with those functions do not fit the data.

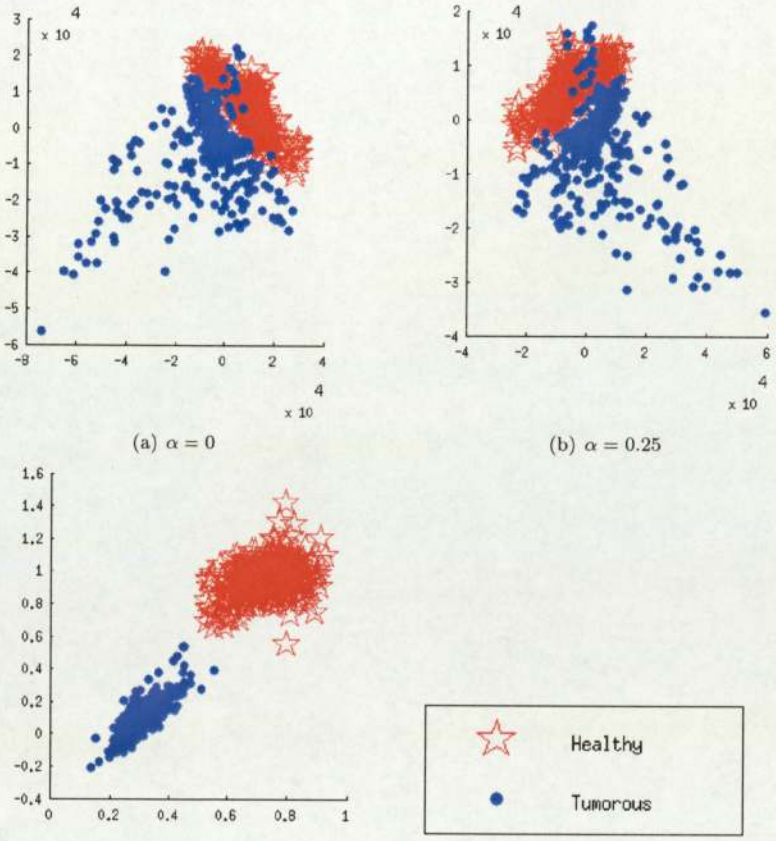
All discrimination cases considered show very strong overlap between the classes considered when using MRI image intensities data. With peak integration data and the combination of peak integration values and MRI data, most of the discrimination cases still show very strong overlap between the classes considered. With α taken equal to 1 (see appendix A.1 on page 98, Tipping, 1996 for details on NeuroScale as well as the use of α), some classes appear to be separable:

- grade 2 gliomas versus grade 3 gliomas
- astrocytoma diffuse grade 2 versus other grade 2
- oligoastrocytoma grade 2 versus other grade 2
- oligoastrocytoma grade 3 versus other grade 3
- oligodendroglioma grade 3 versus other grade 3
- glioma versus meningioma (only when using the combination of MRI and MRSI data)

In the case of healthy versus tumorous, low grade glioma versus high grade glioma and low grade tumour versus high grade tumour, there is some overlap between the classes but the classes seem to be rather easily separable.

We can also notice the large spread of the tumorous class in the healthy versus tumorous case, which is consistent with the large diversity of tumours represented in this class.

The high grade glioma class (and the high grade tumour class since it is the same class) also shows a large spread in the feature space, which represents the intrinsic heterogeneity of tissues affected by this kind of cancer (mixture of low and high grade tissues and dead tissue).

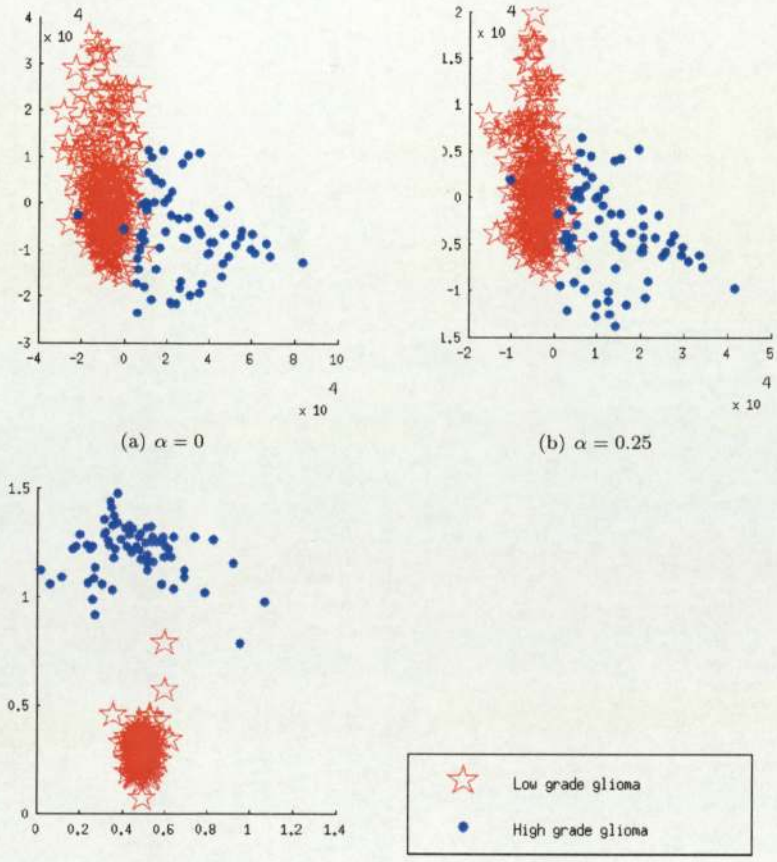


(a) $\alpha = 0$

(b) $\alpha = 0.25$

(c) $\alpha = 1$

Figure 3.9: NeuroScale for the healthy versus tumorous case (combination of MRI and MRSI data).

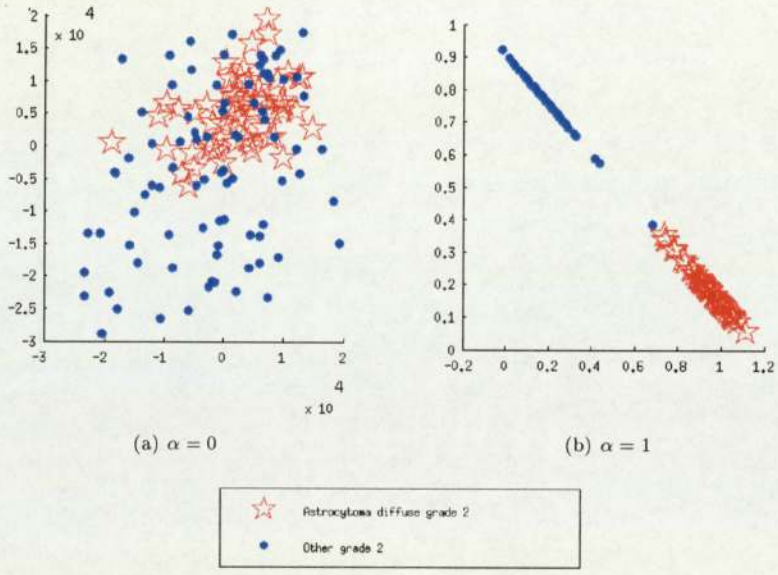


(a) $\alpha = 0$

(b) $\alpha = 0.25$

(c) $\alpha = 1$

Figure 3.10: NeuroScale for the low grade glioma versus high grade glioma case (MRSI data).



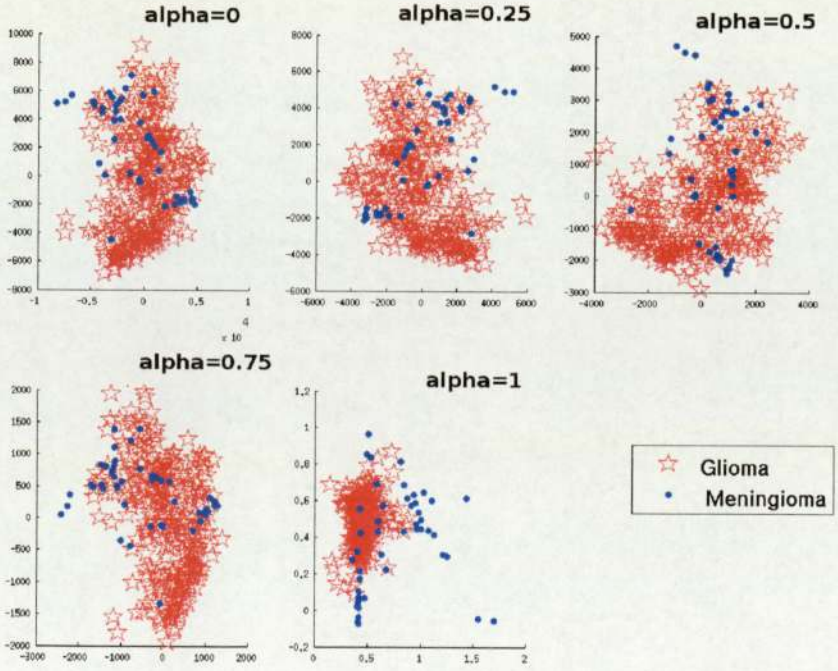
(a) $\alpha = 0$ (b) $\alpha = 1$

Figure 3.11: NeuroScale for the low grade glioma versus high grade glioma case (MRSI and MRI data).

Glioma versus meningioma case

Using only MRI image intensities data, we notice that incorporating class information or not, these two classes do not seem to be separable whereas the two types of cancer they refer to stem from different areas of the brain and should normally be separated straightforwardly by a clinician examining the MRI scans. There can be two hypotheses as to why we are not able to separate these classes correctly just using the MRI image intensities:

- we lose the spatial information given by the MRI scans when we only use MRI image intensities
- the Euclidean distance might not be the appropriate distance to use to discriminate between the two classes.



3.3 Nonlinear Mapping Visualisation in the Assessment Of Brain Tumours

3.3.1 Overview of the method

The difficulty with the type of data we have to deal with is its high dimensionality, which makes inter-patient comparisons and intra-patient analysis difficult. One of the aims of our work is to show how high dimensional data such as this, both inter and intra-patient, can be transformed into a nonlinear topographic feature visualisation process.

A standard dimension-reduction process such as PCA would distort important structural information since it is based on variance maximisation rather than structure preservation. Similarly, many nonlinear machine learning approaches and pattern recognition techniques such as Sammon maps, multidimensional scaling, principal surfaces and self-organising feature maps, do not provide a functional *transformation* of data, only a map of an existing 'training' set. A generative model approach such as GTM would not as well be suitable for our purpose as it is more concerned about modelling the distribution of the data and its density rather than preserving the relationships between data points though these relationships could be indirectly derived with the additional information provided by magnification factors (see Svensen, 1998). And as stated in Svensen, 1998, though an "important feature of the GTM is that, if the mapping from the latent space to the data space is taken to be smooth, the topographic ordering in the latent space will be preserved on the manifold in the data space [....] this does not imply that the GTM is guaranteed to reveal any topographic ordering present in the data". The visualisation method we are using to "overlay" the chemical information extracted from the MRS data and the anatomical information found in the MRI scans relies on pairwise relationships and can be decomposed into the following 6 steps:

Step 1. Our training set consists of the integrated peak values dataset (therefore 669 data points of 10 variables corresponding to voxels belonging to our 25 patients and 4 healthy volunteers). Each of these data points is labelled according to the class it belongs to, healthy or tumorous. We train a NeuroScale model on this data using as many RBF functions for the model as the number of data points of the training set and setting the model output dimension to 2. The training data is then projected by forward propagation through the RBF network obtained with the previous training. This results in a projected dataset of 669 data points in a 2 dimensional space, each data point being labelled as healthy or tumorous.

Step 2. The median of the distribution of the healthy class data points is determined using the output of the NeuroScale model. This median serves as a *reference point* for the computation of distances.

Step 3. The relative distances (computed in the visualisation space) of each of the data

points of the training set (using the NeuroScale model output) to the median of the healthy class is determined, and the maximum distance *cmax* and minimum distance *cmin* is noted for scaling purposes.

Step 4. We project the integrated peak values of each of the patients we have selected by forward propagation through the RBF network obtained with the previous Neuroscale training. The output consists of 2 coordinate variables and of a number of data points equal to the sum of the number of data points for all patients.

Step 5. We compute the distances (computed in the visualisation space) between the patients' data points obtained through forward propagation and the reference point.

Step 6. A single colourmap (containing as many colours as the sum of the number of data points for all patients) is used for all patients to allow inter-patient comparison. For the purposes of the figures in this thesis, this colourmap is split into three parts:

- a part using a scale of shades of green and dark blue corresponding to the values of distances between patients' data points and reference point below *cmin*,
- a part using a scale of shades of cyan and pink corresponding to the values of distances between patients' data points and reference point above *cmax*,
- a part using a colour scale varying smoothly from black through shades of red, orange, and yellow, to white corresponding to the values of distances between patients' data points and reference point between *cmin* and *cmax*

Each of the distances computed in the previous step is associated with a colour of the colormap described above. Knowing the coordinates of each of the patients' data points, we colour the area in the brain that corresponds to the data point in the MRI scan of the patient to which the data point belongs using the colour associated to its relative distance from the reference point.

We now show some final results of this visualisation process.

3.3.2 Some results of the method

We implemented the previous procedure several times using a random initialization for the NeuroScale model and the plots we obtained are consistent and stable. The stability and consistency is not affected by the choice of distance function, and in this thesis we used both Euclidean and Minkowski order 4 metrics. The different metrics used gave different scales of contrast on the map.

The bigger the distance between a given voxel and the reference point (the median of the healthy class), the brighter the shade with which it is coloured. Therefore, as the distance of a voxel from the reference point in the NeuroScale map represents a deviation from "normality" (in this case normality representing healthy), the brighter the colouring of the voxel in the MRI

space, the farther from "normality" the voxel is in chemical similarity space. In that context, we can easily see that in the case of a healthy patient, the obtained voxel colour pattern forms an almost homogeneous red-brown area and that there should be no large area of neighbouring coloured voxels in bright shades (yellow-white), which means that the voxels of the healthy patient are, as expected, close to the reference point (i.e likely to be healthy). In the case of a patient with a brain tumour, the affected areas appear as spread continuums of coloured voxels in brighter yellow-white shades, the size of the area depending on the size of the tumorous area. The brightest-coloured zones might be the area from which the tumour originates as their colouring means they represent the voxels located the farthest from the reference point and therefore should be the voxels of highest grade in the tumorous area and thus very likely to be the ones in which the cancer has had more time to develop.

The voxels coloured in shades of blue-green and shades of magenta-cyan are outliers in the NeuroScale plot. The case where these outliers are coloured in blue-green means that their distance to the reference point therefore to normality is smaller than the minimum distance between the reference point and a voxel belonging to the healthy class, therefore these outliers are very likely to be healthy voxels. In the case where the outliers are coloured in magenta-cyan, the distance between the outliers and the reference point is bigger than the maximum distance between the reference point and the voxels of the training set, suggesting that these outliers are most probably not healthy.

Below are the plots obtained using the Minkowski distance:

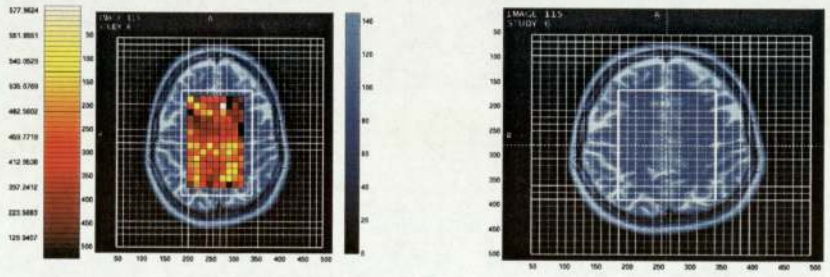


Figure 3.12: Gadolinium-enhanced MRI scan for a healthy volunteer with and without NeuroScale distances mapped on top (Minkowski distance).

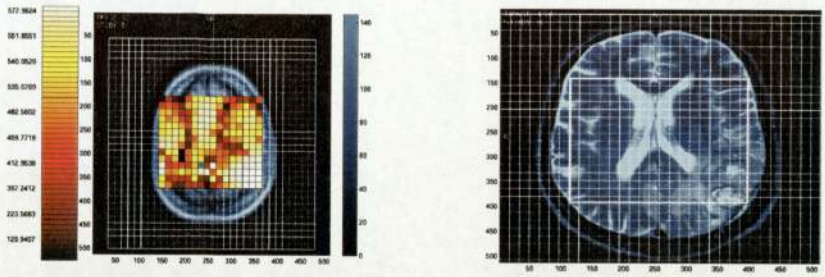


Figure 3.13: Gadolinium-enhanced MRI scan for a glioblastoma patient with and without NeuroScale distances mapped on top (Minkowski distance).

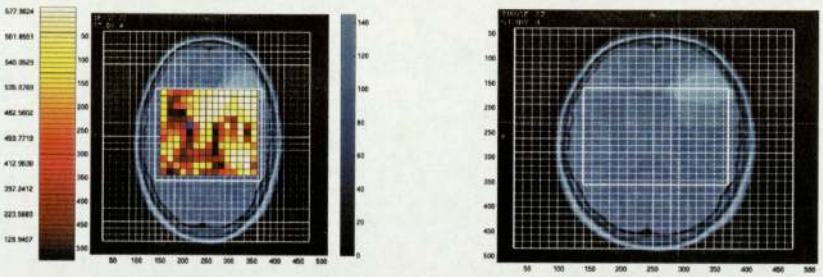


Figure 3.14: Gadolinium-enhanced MRI scan for an oligodendroglioma grade 2 patient with and without NeuroScale distances mapped on top (Minkowski distance).

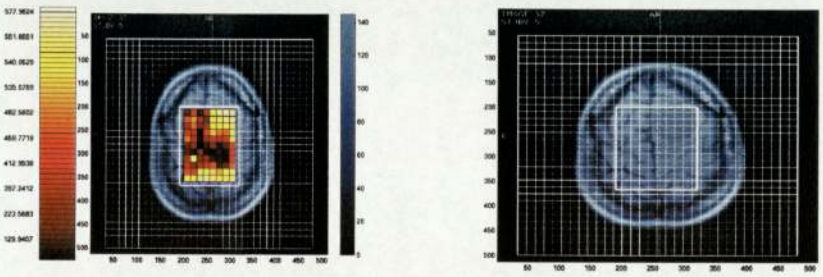


Figure 3.15: Gadolinium-enhanced MRI scan for a meningioma patient with and without NeuroScale distances mapped on top (Minkowski distance).

The following plots have been obtained using the Euclidean distance:

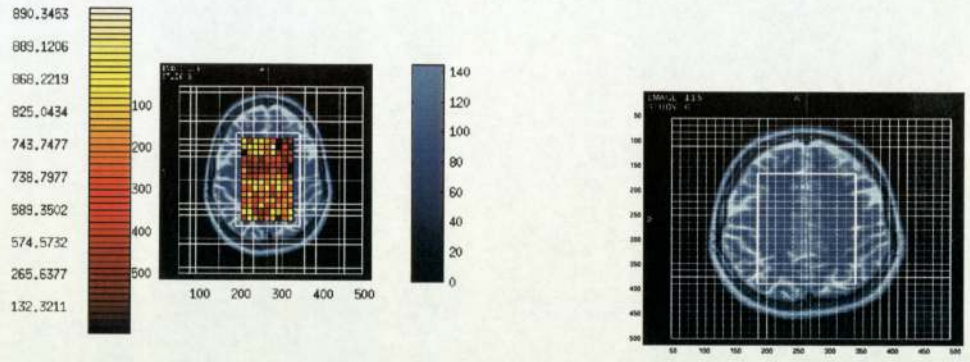


Figure 3.16: Gadolinium-enhanced MRI scan for a healthy volunteer with and without NeuroScale distances mapped on top (Euclidean distance).

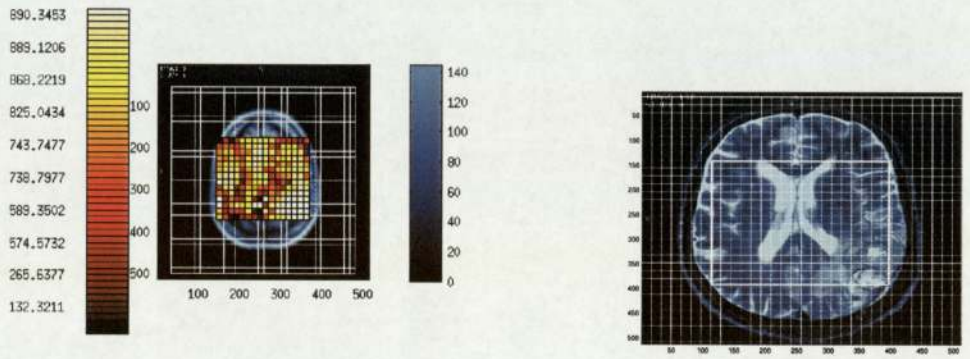


Figure 3.17: Gadolinium-enhanced MRI scan for a glioblastoma patient with and without NeuroScale distances mapped on top (Euclidean distance).

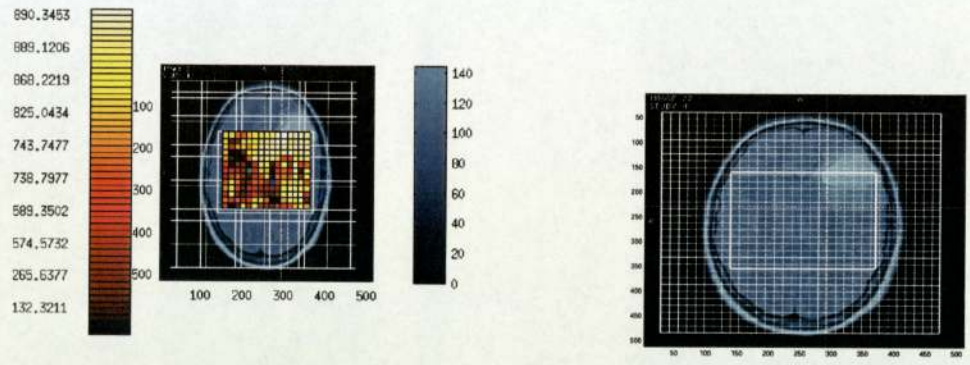


Figure 3.18: Gadolinium-enhanced MRI scan for an oligodendroglioma grade 2 patient with and without NeuroScale distances mapped on top (Euclidean distance).

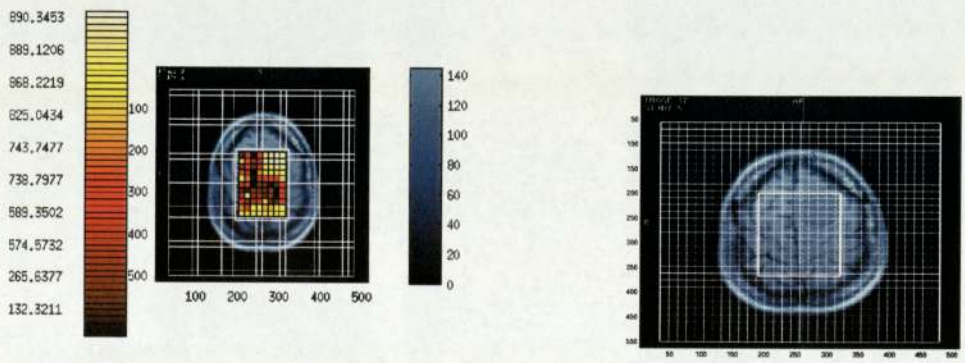


Figure 3.19: Gadolinium-enhanced MRI scan for a meningioma patient with and without NeuroScale distances mapped on top (Euclidean distance).

3.3.3 On the stability of the visualisation

An important issue on the visualisation method used is how stable the method is, in other words for our case how a slight variation on the feature space distances would affect the STRESS function value of the method. Let ϵ be this small variation on the feature space distances. Our STRESS function would be rewritten as

$$E' = \sum_{i}^{N}\sum_{j}^{N}(\delta_{ij}^{*} - ||f(\mathbf{x}_i, \mathbf{w}) - f(\mathbf{x}_j, \mathbf{w}) + \epsilon||)^2. \tag{3.1}$$

.

This is what the equation 3.1 becomes after development:

$$\begin{aligned} E' \quad = \quad & \sum_{i}^{N}\sum_{j}^{N}\delta_{ij}^{*2} - 2\sum_{i}^{N}\sum_{j}^{N}\delta_{ij}^{*}||f(\mathbf{x}_i, \mathbf{w}) - f(\mathbf{x}_j, \mathbf{w}) + \epsilon|| \\ & +.\sum_{i}^{N}\sum_{j}^{N}||f(\mathbf{x}_i, \mathbf{w}) - f(\mathbf{x}_j, \mathbf{w}) + \epsilon||^2 \end{aligned} \tag{3.2}$$

.

If we apply the triangular inequality on the second term of 3.2 and develop the third term then:

$$\begin{aligned} E' \quad \geq \quad & \sum_{i}^{N}\sum_{j}^{N}\delta_{ij}^{*2} - 2\sum_{i}^{N}\sum_{j}^{N}\delta_{ij}^{*}||f(\mathbf{x}_i, \mathbf{w}) \\ & -f(\mathbf{x}_j, \mathbf{w})|| - 2||\epsilon||\sum_{i}^{N}\sum_{j}^{N}\delta_{ij}^{*} \\ & +N^2||\epsilon||^2 + \sum_{i}^{N}\sum_{j}^{N}||f(\mathbf{x}_i, \mathbf{w}) - f(\mathbf{x}_j, \mathbf{w})||^2 \\ & +2||\epsilon||\sum_{i}^{N}\sum_{j}^{N}||f(\mathbf{x}_i, \mathbf{w}) - f(\mathbf{x}_j, \mathbf{w})|| \end{aligned}$$

If we take $\delta = \max_{i,j}\delta_{ij}^{*}$ and $f = \min_{i,j}||f(\mathbf{x}_i, \mathbf{w}) - f(\mathbf{x}_j, \mathbf{w})||$
then we have this lower bound for E':

$E' \geq E - 2N^2||\epsilon||\delta + N^2||\epsilon||^2 + 2||\epsilon||N^2 f \geq E + 2N^2||\epsilon||(f - \delta) + N^2||\epsilon||^2$
where $E = \sum_{i}^{N}\sum_{j}^{N}(\delta_{ij}^{*} - ||f(\mathbf{x}_i, \mathbf{w}) - f(\mathbf{x}_j, \mathbf{w})||)^2$..

In a similar way we find

$E' \leq \sum_i^N \sum_j^N \delta_{ij}^{*2} - 2\sum_i^N \sum_j^N \delta^* ||f + \epsilon|| + N^2 ||f^* + \epsilon||^2 \leq N^2\delta^2 - 2N^2\delta^* ||f + \epsilon|| + n^2 ||f^* + \epsilon||^2$
where $\delta^* = \min_{i,j} \delta_{ij}^*$ and $f^* = \max_{i,j} ||f(\mathbf{x}_i, \mathbf{w}) - f(\mathbf{x}_j, \mathbf{w})||$.

In the upper bound, the dominant term is $n^2\delta$ as the other two terms are of the same order and compensate each other. It shows that the upper error bound would increase with the size of the dataset as well as the spread of the data in the data space.

3.3.4 Conclusion

Using the NeuroScale architecture for nonlinear topographic visualisation on peak integration values data (from MRS), we have been able to map the chemical information from the MRS onto MRI scans in a consistent and stable way. Comparing the colouring of the voxels for which the chemical information from MRS has been mapped and the shades of the neighbouring areas of the MRI scans allows for a rather straightforward interpretation of the scans we obtain through our visualisation technique and therefore for a rather easy detection of potential problems in some areas of the brain. We note that in addition to the actual tumour locations in the brain, there is evidence in some of the overlay scans that apparently 'normal' cells outside the cancerous regions are also showing signs of 'abnormal' chemical behaviour. Evidence of this is given in comparisons of the original spectra obtained from the relevant regions as shown in figure 3.20 on the next page (for the patient with oligodendroglioma grade 2 whose scans are shown in subsection 3.3.2).

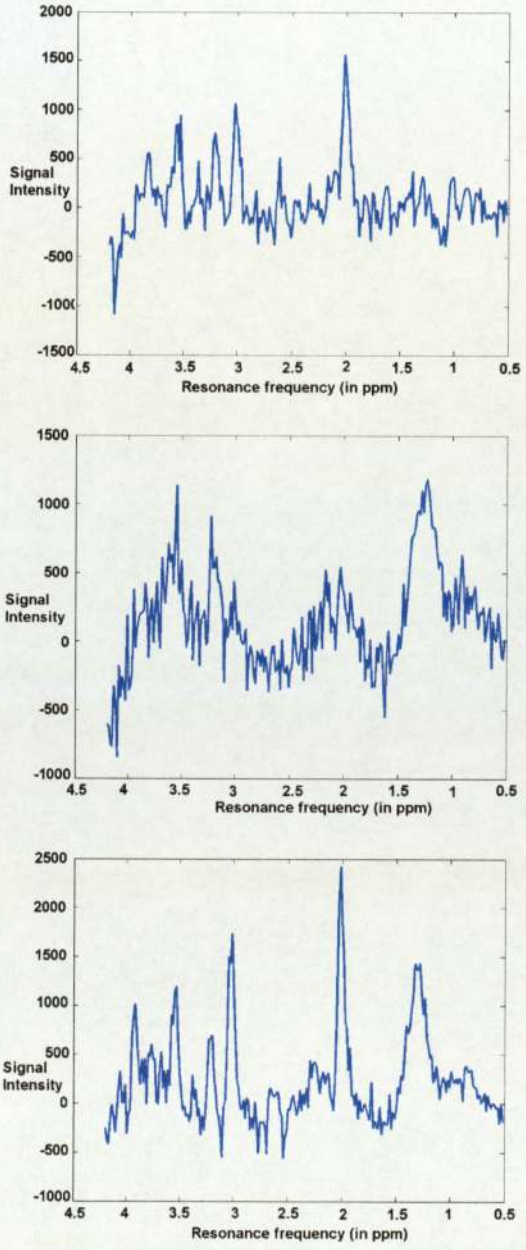


Figure 3.20: Extracted spectra from regions of high colour indicated from the Neuroscale model showing examples of normal (top), cancerous (middle) and 'anomalous' (bottom) cells.

This figure shows spectra of normal, cancerous and anomalous cell regions. It can be seen that the anomalous region spectra shows the strong N-acetylaspartate peak at 2.02 ppm of the normal cell, but is starting to show a large lactate peak representative of the increased tumour metabolism region. Clearly this aspect needs further study. It could, for example, be an early indication of pre-cancerous cell development, or it could possibly be a modified normal cell response to the inter-cell communication processes in which cancerous cells are communicating their modified state to neighbours. Depending on the communication pathways of genetic information, some of these modified response 'normal' cells could be significantly distant from the site of the actual tumour. This of course is currently speculative and requires a more extensive study. Hence these hypotheses should be considered as speculative, awaiting further corroborative evidence.

3.4 Some remarks on normalization and metrics

3.4.1 Normalization

When variations in experimental conditions explain the the quantitative differences of measured values observed for the studied samples, normalisation of data should be performed. After normalisation, the measured data should be adjusted in a way that guarantees that subsequent comparisons of the data samples only uncover biological differences relevant for the research problem addressed. The main assumptions of normalisation are:

- the existence of a functional coherence between a true underlying biological difference characteristic of the biological process studied and the values measured to assess this underlying biological difference. This functional coherence should follow an unequivocal function.
- the presence of a property within the studied samples which is unchanged

In particular, normalisation should be performed to reduce inter-patient and intra-patient variability so as to get a handle on the true underlying biological process (in our case the development of a cancerous process). The intra-patient variability is partly taken care of by the pre-processing steps described in paragraph 1.4.1 on page 26 as the MRSI spectra and the MRI images of each patient are aligned to correct artifacts such as motion artifacts (patient moving during or between the acquisition of MR images and MRSI spectra, physiological movements such as breathing,...), incorrect selection of the brain region for which an scan is made. The pre-processing steps described in paragraph 1.4.1 on page 26 also correct for noise due to the MR equipment (eddy current correction, phase correction, correction for magnetic field inhomogeneities,...) or due to interfering signals (baseline correction, residual water and fat signals removal,...). One of the only sources of intra-patient variability remaining after pre-processing is due to the natural difference between different tissues' chemical and anatomical properties. Therefore, normalization should in our case be mainly used to reduce inter-patient variability. The first idea to achieve this is to normalize by a biological constant supposed to be characteristic for each individual. One such constant would be the creatine level supposed to be constant in different regions of the brain and pathological processes for a given patient. But this creatine level is suspected to vary with

a tumorous process and would not therefore be suitable to perform the normalization. Possible normalization methods could be (our data being contained in matrices where the columns represent the variables and the rows the data points):

- subtracting the mean of the column from each element of the column and then dividing each element of the column by the standard deviation of the column
- subtracting the median of the column from each element of the column and then dividing each element of the column by the mean absolute deviation of the column
- dividing each element of the matrix of data by a given norm of the matrix of data (L2 norm for example)
- normalizing by CSF that is computing the mean vector of the voxels corresponding to CSF and then dividing each element of the ith-column ($i = 1..n$, n being the number of variables (of columns) of the data matrix) by the ith-coordinate of the mean vector previously calculated

In all what follows, the data used will not be normalized.

3.4.2 Investigating different metrics

The metrics, other than the Euclidean distance, that have been investigated are the Mahalanobis distance and the Minkowski distance (of order 4). The results obtained with Minkowski distance of order 4 are similar to the ones obtained with the Euclidean distance and the STRESS function values obtained with this distance are of the order of $10^{12} - 10^{13}$ whereas the STRESS function values for the Euclidean distance are of the order of $10^{11} - 10^{12}$. As for the Mahalanobis distance, the STRESS function values are much lower than the ones obtained for the other distances as they are of the order of $10^5 - 10^6$. The results with MRI data are the same as the ones obtained with the other distances. An increase of α improves dramatically the separability of the classes: at $\alpha = 0$, all classes show very strong overlap whereas starting from $\alpha = 0.5$, all classes appear easily separable.

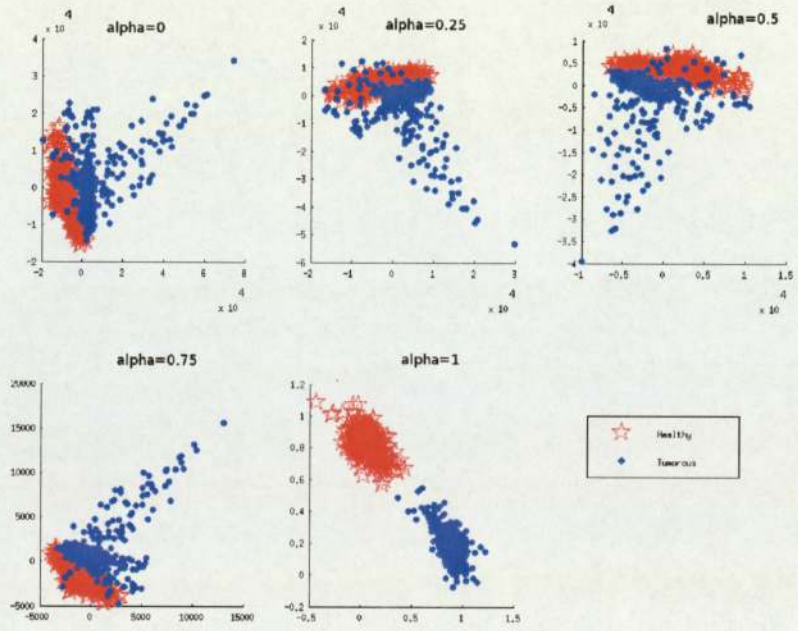


Figure 3.21: NeuroScale of healthy versus tumorous using Minkowski distance of order 4 and the combination of MRI and MRSI data.

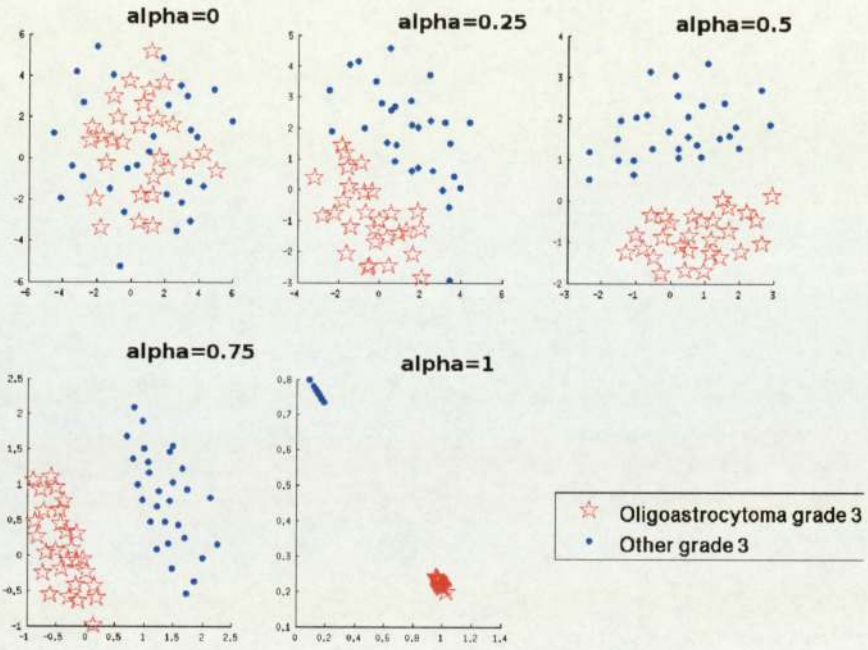


Figure 3.22: NeuroScale of oligoastrocytoma grade 3 versus other grade 3 using Mahalanobis distance and the combination of MRI and MRSI data.

Chapter 4

Classification for the assessment of brain tumours

4.1 Neural Networks Classifiers

4.1.1 Methods

All the classifiers have been used on non-normalized data. Their performance has been tested on MRI image intensities data, peak integration values data as well as the combination of MRI image intensities data and peak integration values data.

The classifiers used are RBF classifiers and MLP classifiers trained in a standard way. RBF classifiers are tested using holdout and leave-one-out procedures. MLP classifiers are tested using holdout and leave-one-out procedures.

The dataset used to construct the classifiers is the part of the training set (described in section 1.4 on page 25 and subsection 1.4.2 on page 27) corresponding to the classes to discriminate. This dataset is split into two parts: a training set and a test set. When using leave-one-out cross-validation, the test set is one of the data points of the dataset used to construct the classifier while the rest of this dataset is the training set used to construct the classifier. When using holdout, the training set is composed of a randomly chosen half of the dataset used to construct the classifier and the test set is the rest of this dataset. [1]

4.1.2 Results

ROC curve analysis

We evaluate the performances of the classifiers by comparing the AUC values obtained with each of the classifiers (for more details on ROC curves and AUC see appendix B on page 100). Our preliminary results on MLP classifiers give very poor performances (AUC values below 0.5). This would need further investigation but for our present aim, highlighting the added value of topographic mapping techniques (especially NeuroScale) with respect to classification approaches, we will not be considering those results and will only be considering the classifiers that give the best AUC performances, in our case the RBF classifiers. The complete results of the classification experiments are shown in appendix D on page 104.

[1] The way of the data is split does not allow the assessment of inter-patient generalisation

The following tables show the minimum and maximum AUC values obtained for each of the classification considered and for each kind of data used.

Table 4.1: Performance of classifiers using peak integration values data

	RBF with holdout	RBF with Leave-one-out
Healthy versus tumorous	0.7445-0.9718	0.8683-0.9814
Glioma versus meningioma	0.9461	0.9814
Low Grade Glioma versus High Grade Glioma	0.9279-0.9944	0.9586-0.9955
Low Grade Tumour versus High Grade Tumour	0.9567-0.9828	0.9673-0.9872
Glioma Grade 2 versus Glioma Grade 3	0.6529-0.9278	0.8072-0.9233
Astrocytoma Grade 2 versus other grade 2	0.6355-0.7987	0.7547-0.822
Oligoastrocytoma Grade 2 versus other grade 2	0-0.7656	0-0.886
Oligoastrocytoma Grade 3 versus other grade 3	0.4809-1	0.7204-1
Oligodendroglioma Grade 3 versus other grade 3	0.482-0.6601	0.605-0.9594

Table 4.2: Performance of classifiers using MRI image intensities data

	RBF with holdout	RBF with Leave-one-out
Healthy versus tumorous	0.7273-0.7756	0.7733-0.7929
Glioma versus meningioma	0.8464-0.8562	0.864-0.8843
Low Grade Glioma versus High Grade Glioma	0.7418-0.8739	0.8735-0.8983
Low Grade Tumour versus High Grade Tumour	0.7736-0.8529	0.7449-0.8891
Glioma Grade 2 versus Glioma Grade 3	0.5402-0.7824	0.5338-0.79
Astrocytoma Grade 2 versus other grade 2	0.2315-0.604	0.3961-0.6733
Oligoastrocytoma Grade 2 versus other grade 2	0.152-0.5351	0.1318-0.498
Oligoastrocytoma Grade 3 versus other grade 3	0.3229-0.6354	0.3793-0.7081
Oligodendroglioma Grade 3 versus other grade 3	0.2614-0.4003	0.2406-0.5331

Table 4.3: Performance of classifiers using the combination of peak integration values data and MRI image intensities data

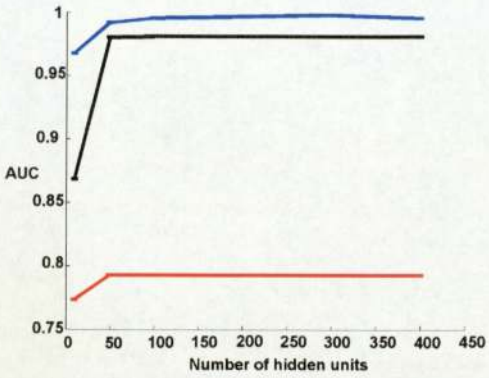
	RBF with holdout	RBF with Leave-one-out
Healthy versus tumorous	0.6571-0.9969	0.968-0.9986
Glioma versus meningioma	0.9726-0.9926	0.9722-0.9964
Low Grade Glioma versus High Grade Glioma	0.9477-0.9984	0.95-0.9956
Low Grade Tumour versus High Grade Tumour	0.9392-0.9925	0.9572-0.9948
Glioma Grade 2 versus Glioma Grade 3	0.7512-0.8858	0.8667-0.9462
Astrocytoma Grade 2 versus other grade 2	0.5285-0.8114	0.6355-0.805
Oligoastrocytoma Grade 2 versus other grade 2	0.1896-0.8043	0.1298-0.8422
Oligoastrocytoma Grade 3 versus other grade 3	0.5486-0.7361	0.6502-1
Oligodendroglioma Grade 3 versus other grade 3	0.4265-0.7157	0.6306-0.9988

As previously outlined in Simonetti, 2004, Devos, 2005, Devos et al., 2005, Lukas, 2003, the best classification results are, in most cases, obtained using the combination of MRI image intensities and peak integration data, the worst results being obtained when using MRI image intensities only (this result is shown in figure 4.1 on the following page in subsection 4.1.2 on page 60).

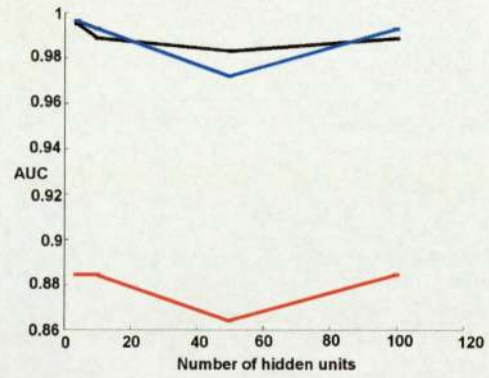
When using the combination of MRI and peak integration values, the RBFs perform very well on most of the classification cases (maximal values of the AUC greater than 0.9). However in some cases, the performances of the classifiers are average to bad:

- grade 2 glioma versus grade 3 glioma
- astrocytoma diffuse grade 2 versus other grade 2
- oligoastrocytoma grade 2 versus other grade 2
- oligoastrocytoma grade 3 versus other grade 3
- oligodendroglioma grade 3 versus other grade 3

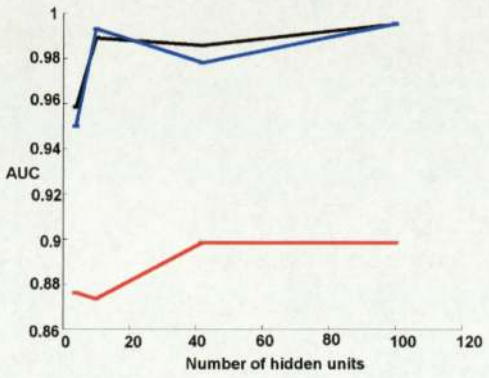
The classifiers' bad performance in the last four cases might be due to the small sample sizes for these particular classes (60 samples of astrocytoma diffuse grade 2, 45 samples of oligoastrocytoma grade 2, 28 samples of oligoastrocytoma grade 3 and 25 samples of oligodendroglioma grade 3).
As for the grade 2 glioma versus grade 3 glioma case, the classifiers' average performance can be explained by the definition of tumour grading: when a voxel is graded as grade 3, it does not mean that all the voxel's tissues are histologically grade 3, it only means that a part of the voxel's tissues, however small it is, is histologically grade 3. In other words, a grade 3 voxel might contain grade 2 tissues as well as grade 3 tissues. This fact makes the grade 2 glioma and grade 3 glioma classes hard to discriminate.
The classifiers' performance improves with the increase of the model complexity (that is the number of hidden units) in most cases as shown in figures 4.1 on the following page, 4.2 on page 65 and 4.3 on page 66 (in subsection 4.1.2 on page 60).



(a) Healthy versus tumorous case

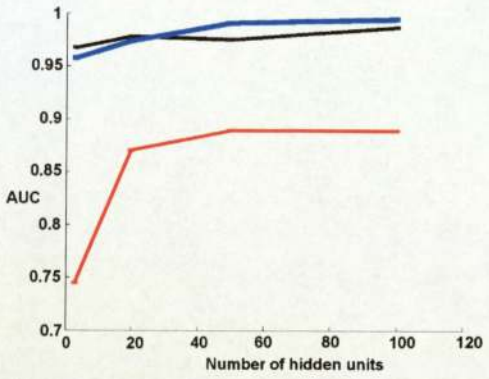


(b) Gliomas versus meningiomas case

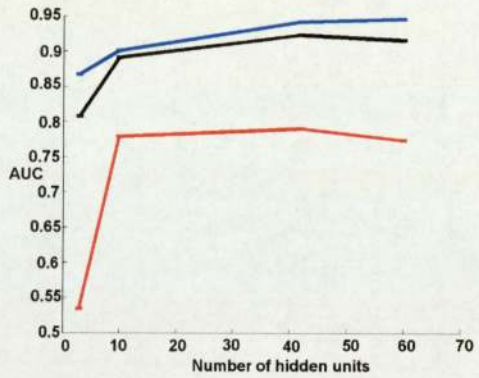


(c) Low grade gliomas versus high grade gliomas case

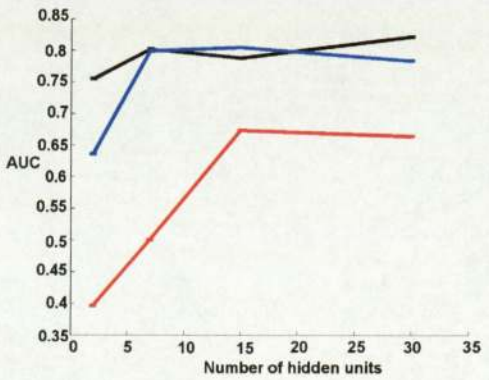
Figure 4.1: RBF classifiers' performances using leave-one-out cross-validation
The black curves are obtained with peak integration values data, the red curves with MRI data and the blue curves with the combination of MRI and peak integration values data.



(a) Low grade tumours versus high grade tumours case

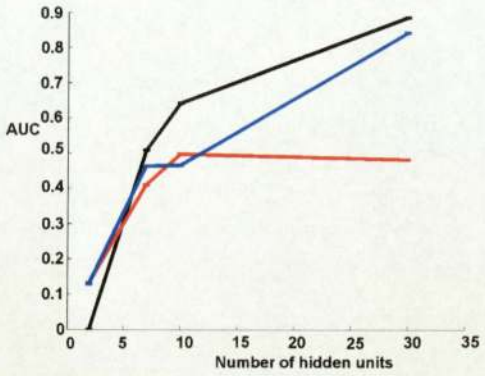


(b) Grade 2 gliomas versus Grade 3 gliomas case

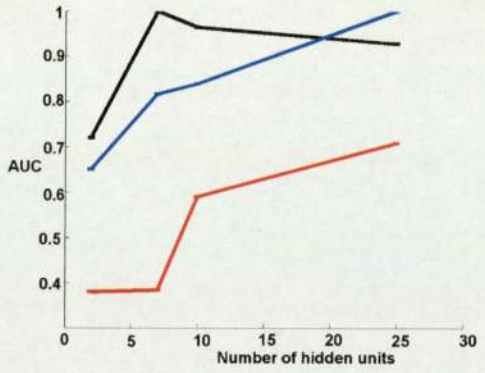


(c) Astrocytoma diffuse grade 2 versus other grade 2 gliomas case

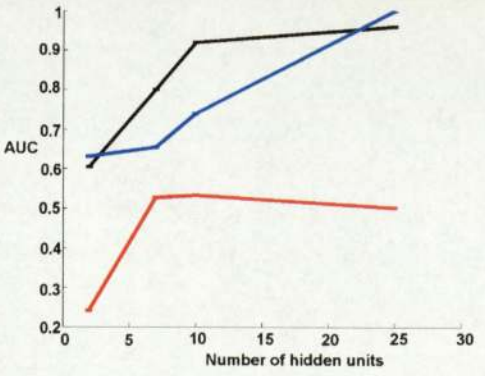
Figure 4.2: RBF classifiers' performances using leave-one-out cross-validation
The black curves are obtained with peak integration values data, the red curves with MRI data and the blue curves with the combination of MRI and peak integration values data.



(a) Oligoastrocytoma grade 2 versus other grade 2 gliomas case



(b) Oligoastrocytoma grade 3 versus other grade 3 gliomas case



(c) Oligodendroglioma grade 3 versus other grade 3 gliomas case

Figure 4.3: RBF classifiers' performances using leave-one-out cross-validation
The black curves are obtained with peak integration values data, the red curves with MRI data and the blue curves with the combination of MRI and peak integration values data.

4.1.3 Mapping the classifier outputs on MRI scans

The process used to map the classifiers outputs for a patient is similar to the one used to overlay NeuroScale distances on MRI scans. This mapping is done as follows:

Step 1. The MRSI data of each patient (taken from the patient-specific data described in section 1.4 on page 25 and subsection 1.4.3 on page 28) is projected by forward propagation through each of the RBF classifiers obtained with the previous training. This results is a projected dataset of as many data points as the sum of the number of acquired MRSI voxels for all the patients considered. This dataset contains two variables : the probability of each of the voxels to be healthy and its probability to be tumorous [2].

Step 2. The variable we are really interested in is each voxel's probability to be tumorous. A single colourmap (containing as many colours as the sum of the number of data points for all patients) is chosen to allow inter-patient comparison. To each value of the variable of interest (probability of being tumorous) is associated a colour of the chosen colourmap. Knowing the coordinates of each of the patients' data points, we colour the area in the brain that corresponds to the data point in the MRI scan of the patient to which the data point belongs using the colour associated to its probability of being tumorous.

The following figures (figures 4.5 on page 69, 4.6 on page 70, 4.8 on page 72 and 4.7 on page 71 (subsection 4.1.3)) correspond to the overlayed scans obtained for the patients for which we have provided NeuroScale distances overlayed scans in section: a patient with glioblastoma multiforme, a patient with meningioma, a healthy volunteer and a patient with oligodendroglioma grade 2. For each of the figures, the classifiers were obtained using leave-one-out cross-validation and a classification threshold of 0.6. The threshold of 0.6 as well as the models with 100, 300 and 400 hidden units are chosen as they are those that give the best tradeoff sensitivity/false positive rate (that is those that maximize the sensitivity while keeping the false positive rate as low as possible) (see the ROC curve in figure 4.4 on the next page (subsection 4.1.3)).

[2]We ensure that the obtained probabilities are in the range [0-1] by applying the softmax function on the outputs of the RBF networks.

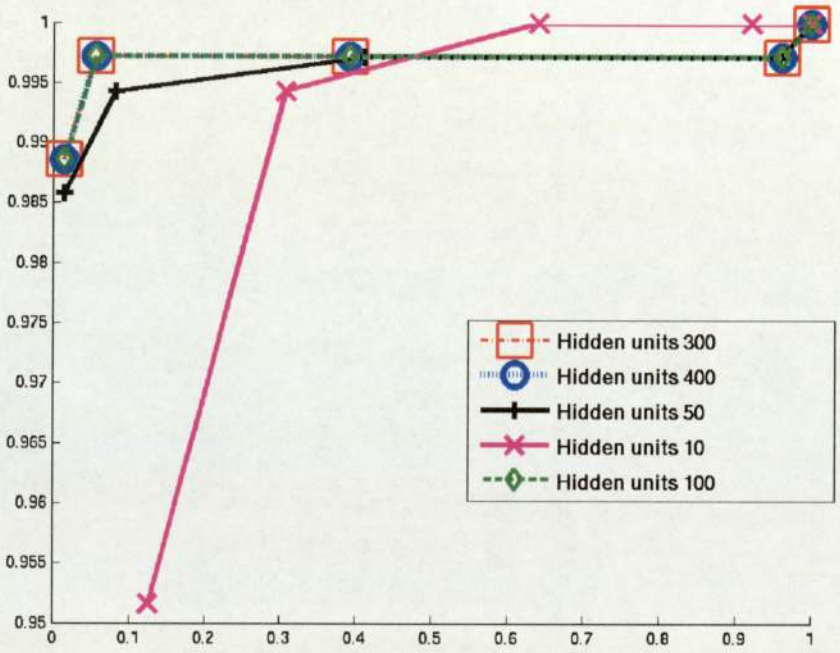
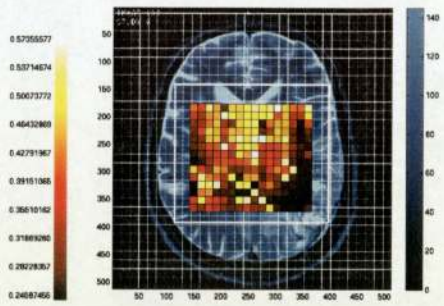
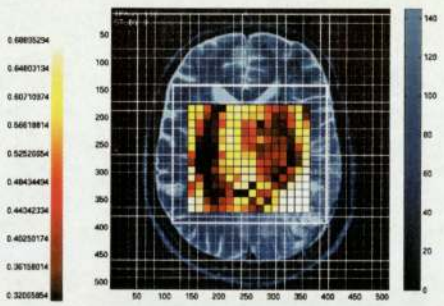


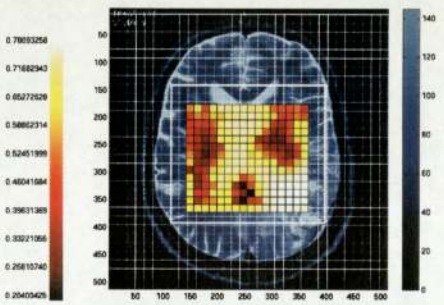
Figure 4.4: ROC curve of the RBF classifiers for healthy versus tumourous (using leave-one-out cross-validation).
For each curve, the first data point (starting from the left) corresponds to the data point for a threshold of 0.5, the second to the data point for a threshold of 0.6, the third to the data point for a threreld of 0.7, the fourth to the data point for a threshold of 0.8 and finally the fifth to the data point for a threshold of 0.9.



(a) Classifier with 100 hidden units
Scale of probabilities 0.2451-0.5736

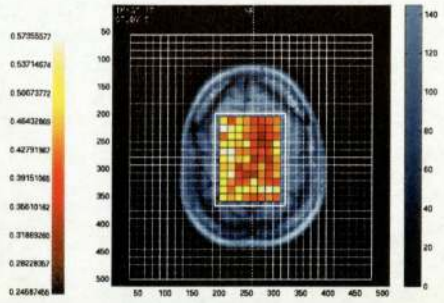


(b) Classifier with 300 hidden units
Scale of probabilities 0.3207-0.6890

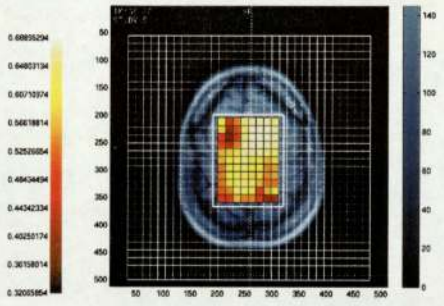


(c) Classifier with 400 hidden units
Scale of probabilities 0.2040-0.7809

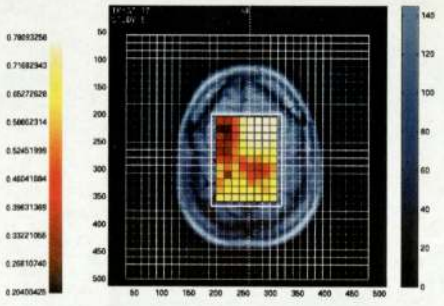
Figure 4.5: MRI scans of a patient with glioblastoma multiforme overlayed with the probabilities of each voxel to belong to the tumorous class (using leave-one-out cross-validation).



(a) Classifier with 100 hidden units
Scale of probabilities 0.2451-0.5736

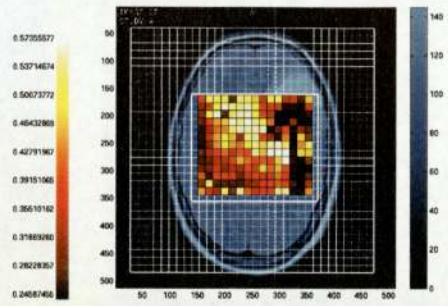


(b) Classifier with 300 hidden units
Scale of probabilities 0.3207-0.6890

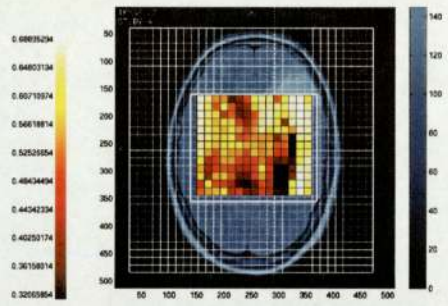


(c) Classifier with 400 hidden units
Scale of probabilities 0.2040-0.7809

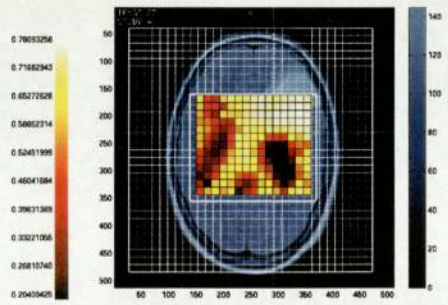
Figure 4.6: MRI scans of a patient with meningioma overlayed with the probabilities of each voxel to belong to the tumorous class (using leave-one-out cross-validation).



(a) Classifier with 100 hidden units
Scale of probabilities 0.2451-0.5736

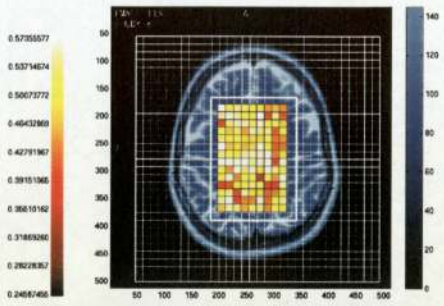


(b) Classifier with 300 hidden units
Scale of probabilities 0.3207-0.6890

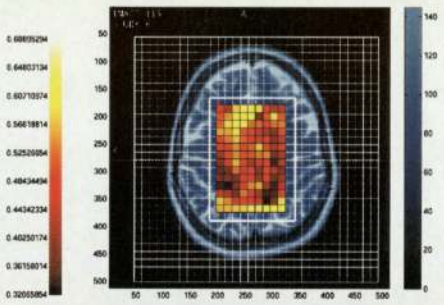


(c) Classifier with 400 hidden units
Scale of probabilities 0.2040-0.7809

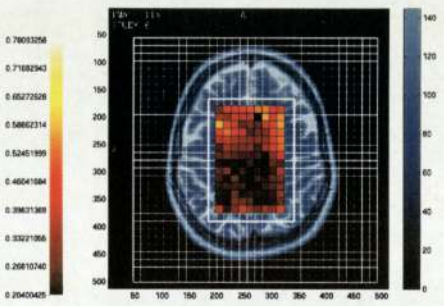
Figure 4.7: MRI scans of a patient with oligodendroglioma grade 2 overlayed with the probabilities of each voxel to belong to the tumorous class (using leave-one-out cross-validation).



(a) Classifier with 100 hidden units
Scale of probabilities 0.2451-0.5736



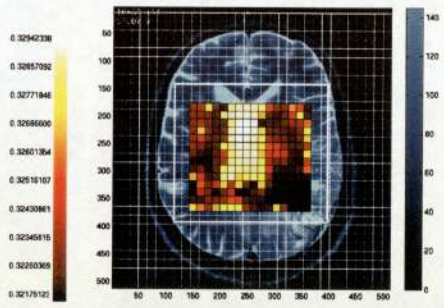
(b) Classifier with 300 hidden units
Scale of probabilities 0.3207-0.6890



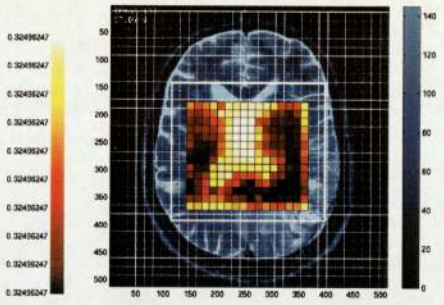
(c) Classifier with 400 hidden units
Scale of probabilities 0.2040-0.7809

Figure 4.8: MRI scans of a healthy volunteer overlayed with the probabilities of each voxel to belong to the tumorous class (using leave-one-out cross-validation).

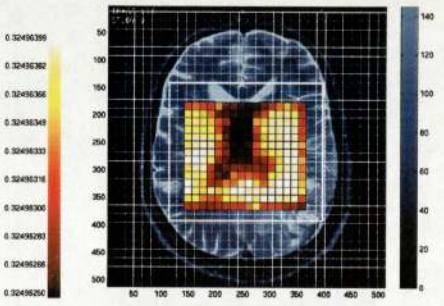
The following figures (figures 4.9 on the next page, 4.10 on page 75, 4.12 on page 77 and 4.11 on page 76) correspond to the overlayed scans obtained for the same patients as before using classifiers obtained using holdout (that is half of the training data randomly taken as training set and the other half as test set) and a classification threshold of 0.6.



(a) Classifier with 100 hidden units
Scale of probabilities 0.3218-0.3294

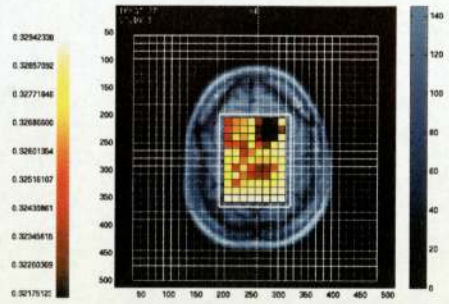


(b) Classifier with 300 hidden units
Scale of probabilities 0.3249624726230-0.3249624726232

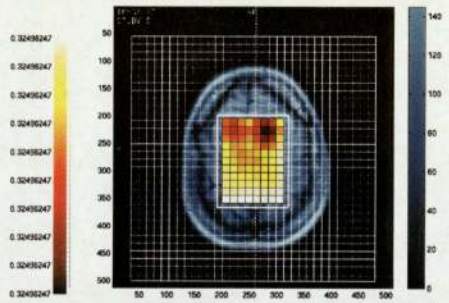


(c) Classifier with 400 hidden units
Scale of probabilities 0.324962-0.324964

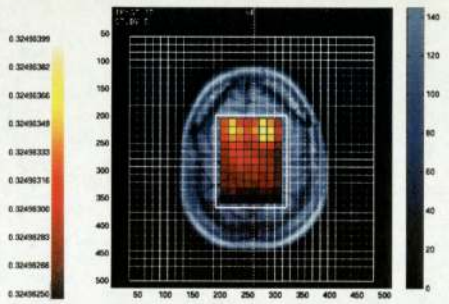
Figure 4.9: MRI scans of a patient with glioblastoma multiforme overlayed with the probabilities of each voxel to belong to the tumorous class (with holdout).



(a) Classifier with 100 hidden units
Scale of probabilities 0.3218-0.3294

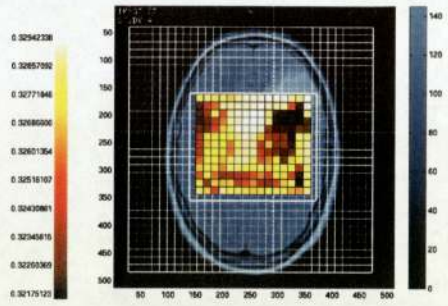


(b) Classifier with 300 hidden units
Scale of probabilities 0.3249624726230-0.3249624726232

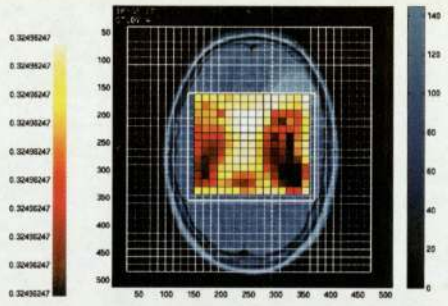


(c) Classifier with 400 hidden units
Scale of probabilities 0.324962-0.324964

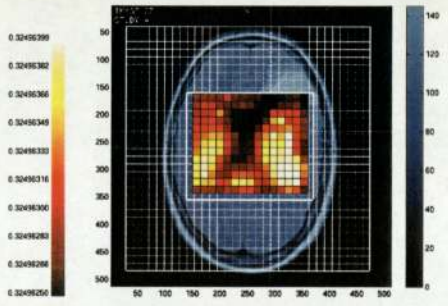
Figure 4.10: MRI scans of a patient with meningioma overlayed with the probabilities of each voxel to belong to the tumorous class (with holdout).



(a) Classifier with 100 hidden units
Scale of probabilities 0.3218-0.3294

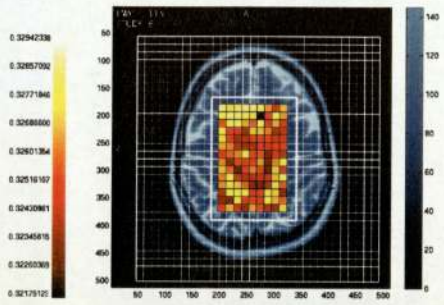


(b) Classifier with 300 hidden units
Scale of probabilities 0.3249624726230-0.3249624726232

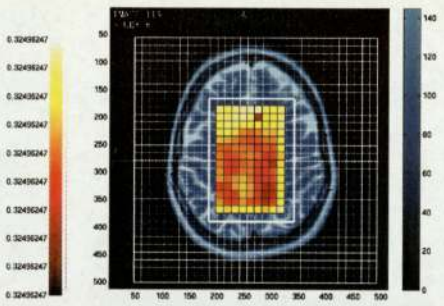


(c) Classifier with 400 hidden units
Scale of probabilities 0.324962-0.324964

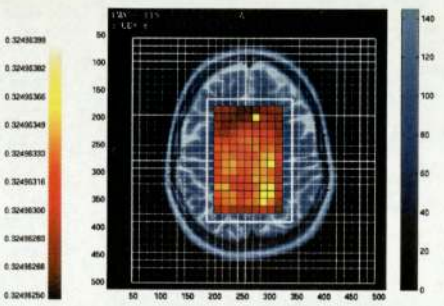
Figure 4.11: MRI scans of a patient with oligodendroglioma grade 2 overlayed with the probabilities of each voxel to belong to the tumorous class (with holdout).



(a) Classifier with 100 hidden units
Scale of probabilities 0.3218-0.3294



(b) Classifier with 300 hidden units
Scale of probabilities 0.3249624726230-0.3249624726232



(c) Classifier with 400 hidden units
Scale of probabilities 0.324962-0.324964

Figure 4.12: MRI scans of a healthy volunteer overlayed with the probabilities of each voxel to belong to the tumorous class (with holdout).

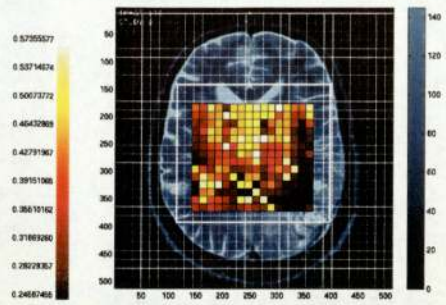
If we look at the patient with glioblastoma multiforme, we can notice that, as in NeuroScale overlayed MRI scans, CSF regions are classified as being tumorous by the classifiers with 300 and 400 hidden units using leave-one-out cross-validation (with a classification threshold of 0.6). This could be the result of CSF carrying tumorous cells (and therefore is classed as tumorous) is the basis of the use of lumbar puncture to diagnose brain tumours, but the concentration of tumourous cells in CSF would probably not be enough to trigger so marked a deviation from normality. This fact could also be explained by there not being an independent CSF class. It shows nevertheless that CSF has very different chemical properties from non-CSF healthy tissue and that CSF cannot be considered as belonging to the healthy class.
The other fact we can notice is low scale of probabilities of some classifiers (for example the classifiers using holdout whose outputs are shown in figures 4.9 on page 74, 4.10 on page 75, 4.12 on the previous page and 4.11 on page 76 (subsection 4.1.3 on page 67)), which would then classify all the patients' data points (even those used in the training and known to be tumorous) as non-tumorous. This could be due to the poor generalization ability of these classifiers. These classifiers having very poor performances should be discarded.

Chapter 5

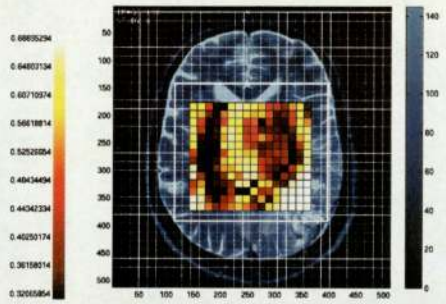
Comparison between topographic mapping results and classification results

5.1 Methods

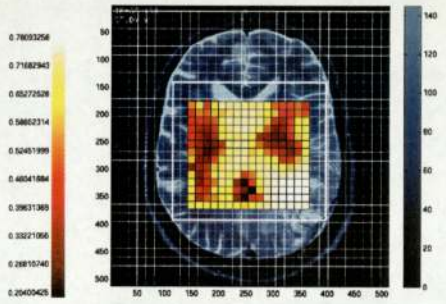
The overlayed MRI scans obtained using the Euclidean distance and using Minkowski distance order 4 being similar, we will rely on the overlayed MRI scans obtained with the Euclidean distance for our comparison of classifiers and use of NeuroScale distances' computations. We will compare those scans with the classifiers' probabilities overlayed MRI scans. The classifiers' probabilities overlayed MRI scans that have been chosen for the comparison are those that "maximize" the sensitivity (that is the true positive rate) while "minimizing" the false positive rate (that is 1-specificity). In our case the classifiers chosen have a sensitivity of 0.9972 and a false positive rate of 0.0566 and correspond to the models with 100, 300 and 400 hidden units and a classification threshold of 0.6 (which means that a voxel whose probability of belonging to the tumorous class is above 0.6 is classified as tumorous).
A first inspection of the probabilities-overlayed MRI scans shows us that the plots we obtain for each of the chosen classifiers are not consistent and stable and give different scales of contrast on the map even though the chosen classifiers give the same classification performance (that is the same sensitivity and specificity) (see figures 5.1 on the following page and 5.2 on page 81 in section 5.1).



(a) Classifier with 100 hidden units
Scale of probabilities 0.2451-0.5736

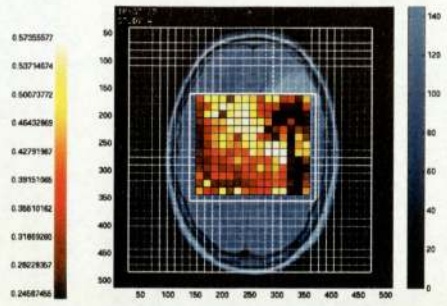


(b) Classifier with 300 hidden units
Scale of probabilities 0.3207-0.6890

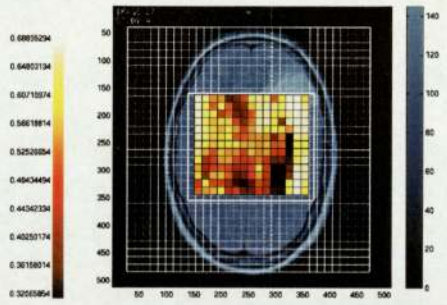


(c) Classifier with 400 hidden units
Scale of probabilities 0.2040-0.7809

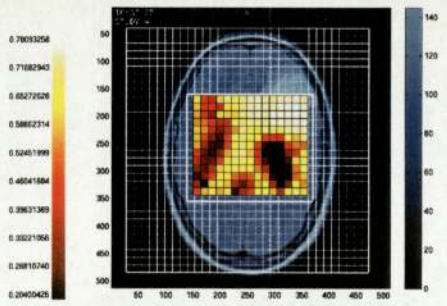
Figure 5.1: MRI scans of a patient with glioblastoma multiforme overlayed with the probabilities of each voxel to belong to the tumorous class.



(a) Classifier with 100 hidden units
Scale of probabilities 0.2451-0.5736



(b) Classifier with 300 hidden units
Scale of probabilities 0.3207-0.6890



(c) Classifier with 400 hidden units
Scale of probabilities 0.2040-0.7809

Figure 5.2: MRI scans of a patient with oligodendroglioma grade 2 overlayed with the probabilities of each voxel to belong to the tumorous class.

To compare the classifiers' outputs and topographic mapping results more thoroughly, we will consider the NeuroScale overlayed MRI scans and probabilities overlayed MRI scans of 2 patients (a patient affected with glioblastoma multiforme and a patient with oligodendroglioma grade 2) and look for the regions where the predictions of the classifiers and those obtained with the computation of NeuroScale distances are different (for example regions where the probability of belonging to the tumorous class is low whereas the distance to the reference point in the NeuroScale feature space is large (regions in bright shades of yellow and white)).
For each of the regions identified, we will be looking at the MRS spectra of the region to determine whether the region might be classified as tumorous or not. In the MRS spectra we will be principally focusing on the following metabolites:

- choline (3.20 ppm)
- lactate (1.33 ppm)
- myoInositol (3.56 ppm)
- creatine (3.02 ppm)
- NAA (2.02 ppm)

Choline is a precursor of acetylcholine (a neurotransmitter) and a component of certain phospholipids. As choline-containing compounds are involved in the synthesis and degradation of cell membranes, high levels of choline-containing compounds indicate increased cellular activity and proliferation among cerebral cells (for example glial cells in the case of gliomas). Therefore an increased choline level, because it indicates an increased cell proliferation and activity, is strongly indicative of a tumorous process.
Healthy cells tend to consume glucose using an aerobic process which doesn't produce lactate. Hence a low level of lactate in healthy cells. An increase in lactate levels shows the consumption of glucose is done through anaerobic pathways and is indicative of tumour metabolism as tumours tend to consume glucose using only anaerobic pathways.
Creatine is a substance which plays an important role in energy metabolism and is phosphorylated (i.e phosphate groups are added to it) into phosphocreatine which is an energy storage compound. Therefore high levels of creatine suggest an increased energy metabolism and would be more consistent with cancerous cells than with healthy cells.
MyoInositol is an isomer of glucose found in cell membrane phospholipids, plasma lipoproteins, and in the nucleus with potential chemopreventive properties, when phosphorylated in phosphatidylinositol. As one of a number of intracellular phosphate compounds, myoInositol is involved in cell signaling and may stimulate tumour cell differentiation. High levels of myoInositol also indicate glial hypertrophy and proliferation implying a cancerous process.
N-acetylaspartate (NAA) is an amino-acid primarily found in neurons. Because of its predominant neuronal localization as well as the strong acetate signal it produces in water-suppressed

proton MRS, NAA is used as an indicator of neuronal viability. Reduced levels of NAA therefore indicate neuronal damage, loss or functional degeneration that may be associated with tumours (among other diseases). High levels of NAA on the contrary indicate healthy cells.
If we look more closely at the classifier using 100 hidden units, we can see that none of the voxels' probabilities of belonging to the tumorous class is above 0.57355577 (for all the patients considered), which means, that as the threshold is set to 0.6, none of the voxels would belong to the tumorous class, if this classifier is used, even those that have been identified as tumorous after histological examination as in the case of the patients with glioblastoma multiforme and oligodendroglioma grade 2. Because of this, we will not consider this classifier for the comparisons with the NeuroScale results.

5.2 Some case studies

5.2.1 Patient with glioblastoma

For this patient we will study 5 voxels classified by the two classifiers considered as healthy (probability of being tumorous between 0.3 and 0.4 for those voxels) as well as 3 other voxels classified as healthy by the classifier with 300 hidden units and represented in bright yellow-white shades in the NeuroScale distances overlayed scan. The first group of voxels considered are the voxels labelled 83, 84 and 85 (see figure 5.3 on the next page in subsection 5.2.1).

The spectrum of voxel 85 (figure 5.4 on page 85 in subsection 5.2.1) plainly shows a decreased level of NAA as well as increased levels of myoInositol and choline, highly suggestive of cancer. Though the spectrum of voxel 84 (figure 5.5 on page 85in subsection 5.2.1) shows a strong peak of NAA, it also exhibits a very high level of choline (almost as high as the NAA peak). In this case, we cannot say the voxel is tumorous, yet it is clear that this voxel has an 'abnormal' chemical behaviour. This behaviour could be due to the fact that the tissues of this voxel are heterogeneous (that is partly healthy and partly tumorous) or to a early stage of cancerous cell development or to a modified normal cell response to the inter-cell communication processes in which cancerous cells are communicating their modified state to neighbours.

The spectrum of voxel 83 (figure 5.6 on page 86in subsection 5.2.1) exhibits the same behaviour as that of voxel 84.

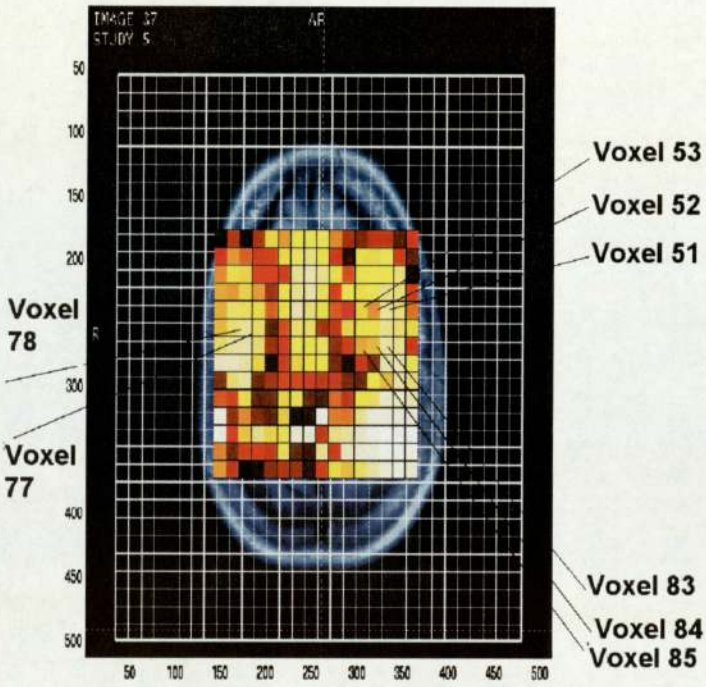


Figure 5.3: Location of the studied voxels for a patient with glioblastoma multiforme (plot overlayed with NeuroScale distances).

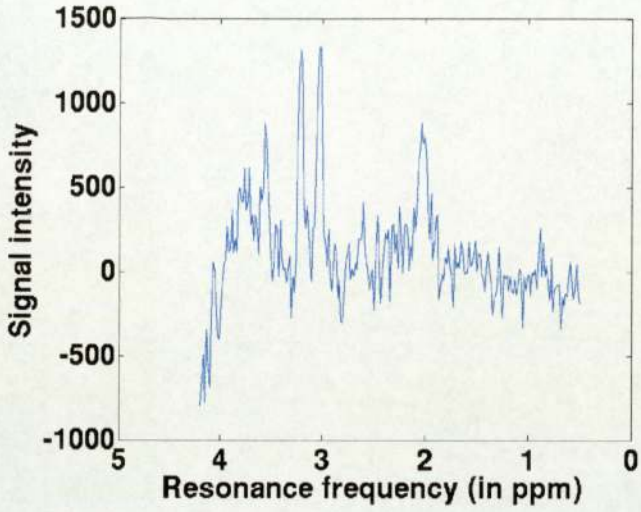


Figure 5.4: Voxel 85 of a patient with glioblastoma multiforme.

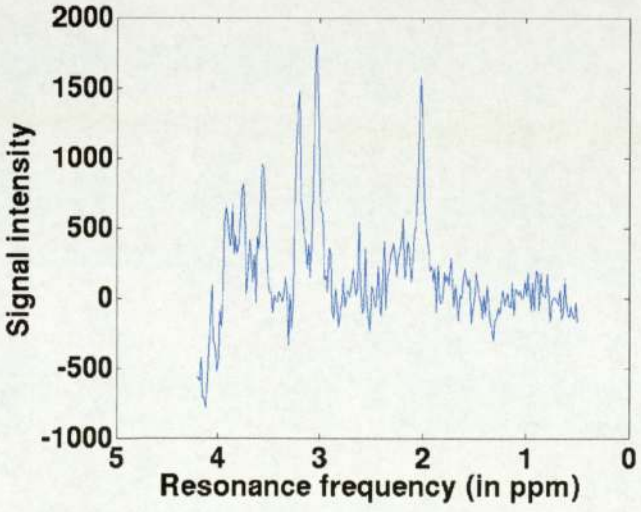


Figure 5.5: Voxel 84 of a patient with glioblastoma multiforme.

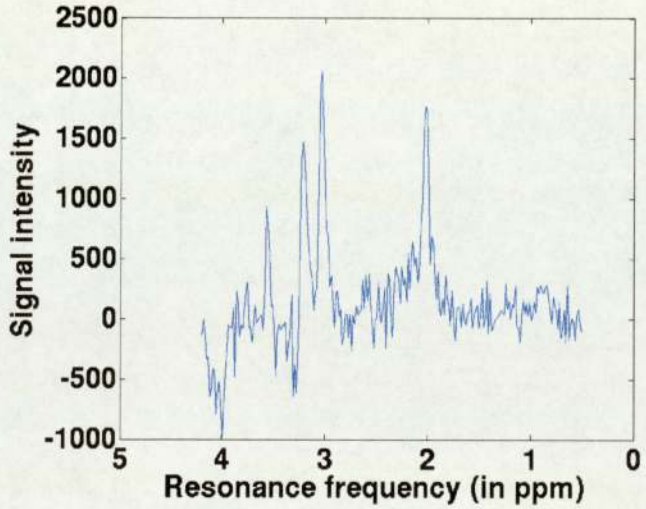


Figure 5.6: Voxel 83 of a patient with glioblastoma multiforme.

For this group of voxels, NeuroScale has been able to identify a voxel misclassified as healthy (voxel 85) by the two considered classifiers and two 'anomalous' voxels (voxels 83 and 84).
The second group of voxels is the group of voxels labelled 51, 52 and 53 (see figure 5.3 on page 84). This group is identified as tumorous by the classifier using 400 hidden units but as healthy by the classifier with 300 hidden units and is represented in shades of yellow in NeuroScale. These voxels (figures 5.7, 5.8 on the next page and 5.9 on the following page in subsection 5.2.1 on page 83) show very strong peaks of choline and very low peaks of NAA. This chemical behaviour is strongly suggestive of cancer. Furthermore, a strong peak of lactate/lipids is visible in the spectrum of voxel 51, which should indicate an advanced stage of cancer development. Therefore for this group of voxels, NeuroScale confirms the classification of the classifier with 400 hidden units, infirms the classification of the classifier with 300 hidden units and is able to identify correctly the tumorous status of the voxels.

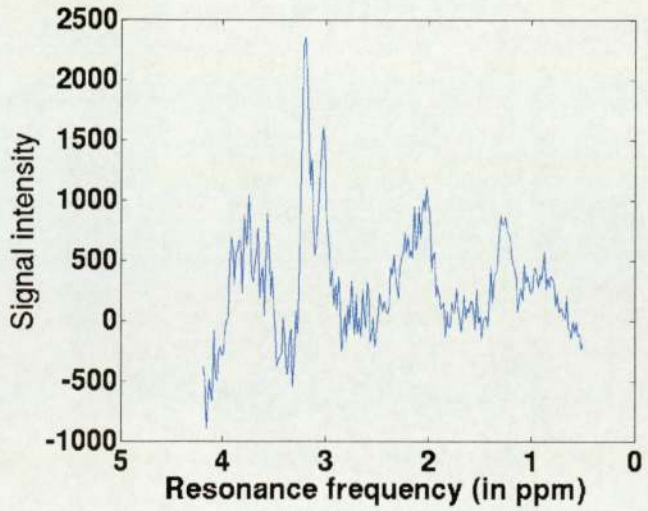


Figure 5.7: Voxel 51 of a patient with glioblastoma multiforme.

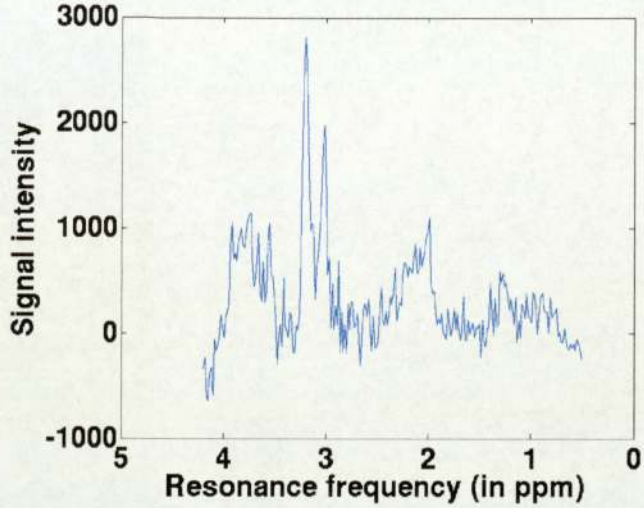


Figure 5.8: Voxel 52 of a patient with glioblastoma multiforme.

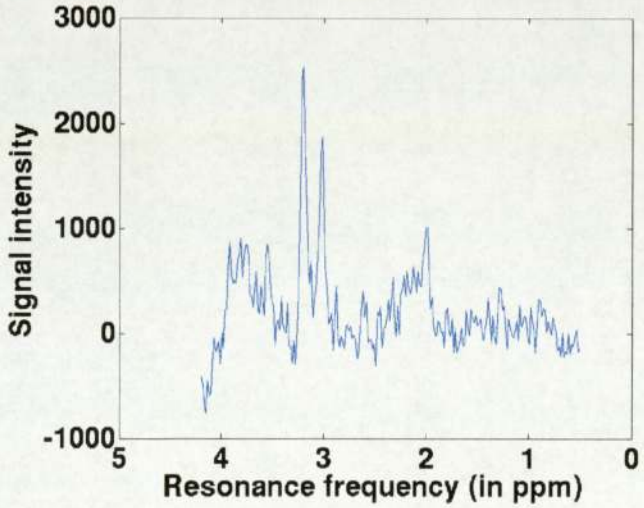


Figure 5.9: Voxel 53 of a patient with glioblastoma multiforme.

The third group of voxels is the group of voxels 77 and 78 (see figure 5.3 on page 84 in subsection 5.2.1 on page 83). The spectra of these voxels (figures 5.2.1 and 5.11 on the next page in subsection 5.2.1 on page 83) exhibit very high peaks of NAA, choline as well as myoInositol, which might indicate an early stage of cancer development or the heterogeneous nature of these voxels' tissues. Here as well, NeuroScale identifies an anomalous chemical behaviour of the voxels, not identified by either of the classifiers considered.

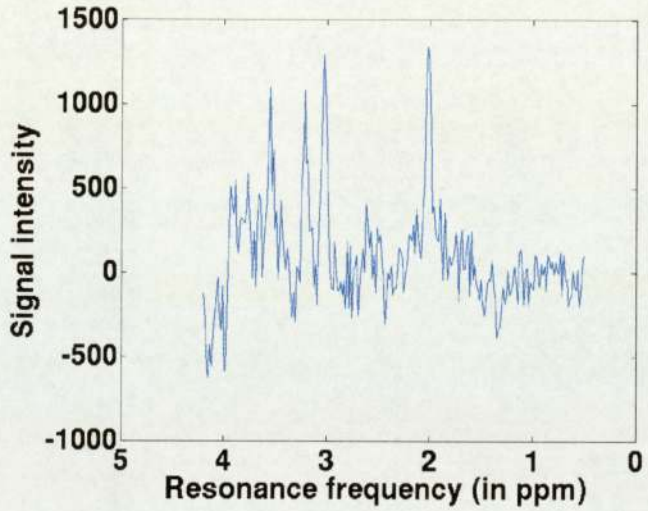


Figure 5.10: Voxel 77 of a patient with glioblastoma multiforme

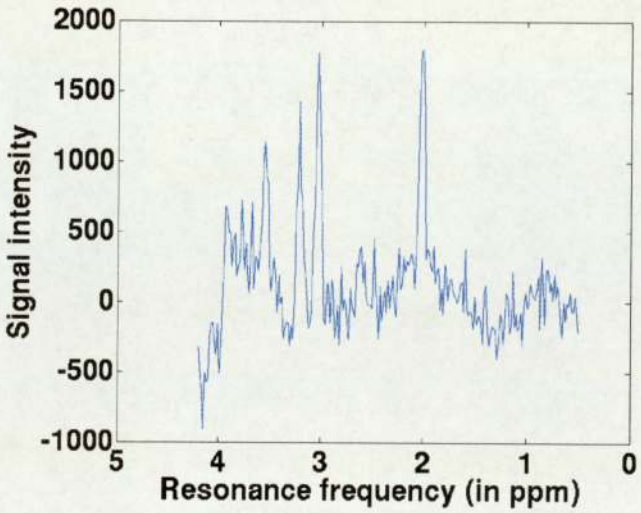


Figure 5.11: Voxel 78 of a patient with glioblastoma multiforme.

5.2.2 Patient with oligodendroglioma grade 2

For this patient, we will consider 3 voxels classified as healthy by the classifier with 400 hidden units with probabilities of being tumorous between 0.2 and 0.3 and as healthy by the classifier with 300 hidden units with probabilities of being tumorous between 0.3 and 0.4 (and whose location is shown in figure 5.12in subsection 5.2.2).

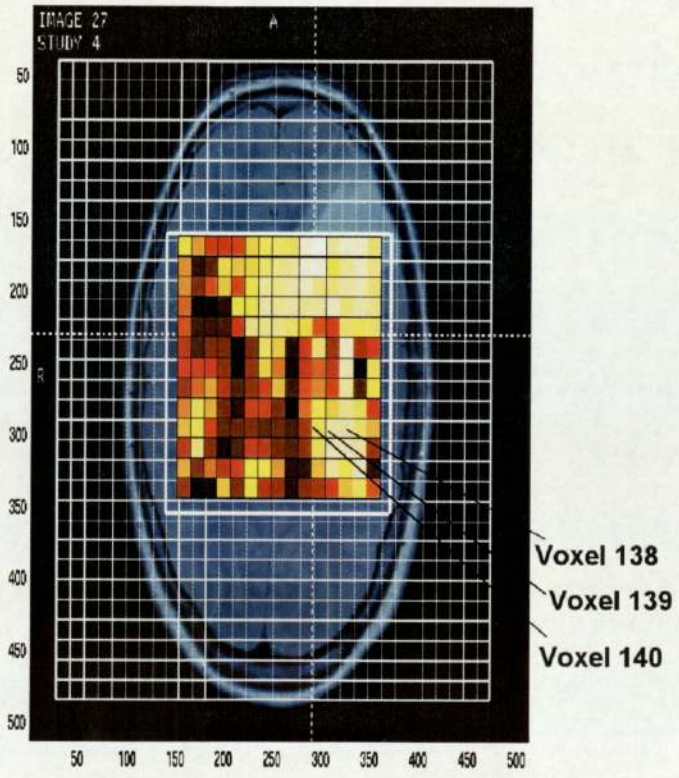


Figure 5.12: Location of the studied voxels for a patient with oligodendroglioma grade 2 (plot with overlayed NeuroScale distances).

These three voxels (voxels 138, 139 and 140) show decreased levels of NAA, very high levels of myoInositol, non negligible levels of lactate and increased levels of choline, thus suggesting a cancerous behaviour. Here NeuroScale helps identify a tumorous behaviour of the voxels, not identified by either of the classifiers.

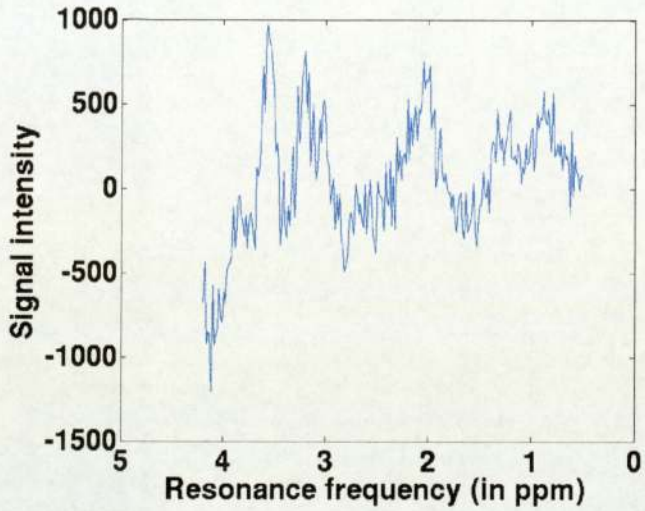


Figure 5.13: Voxel 138 of a patient with oligodendroglioma grade 2.

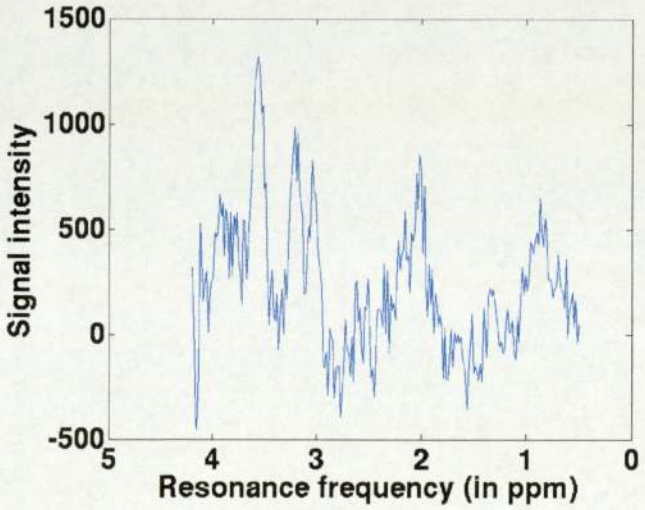


Figure 5.14: Voxel 139 of a patient with oligodendroglioma grade 2.

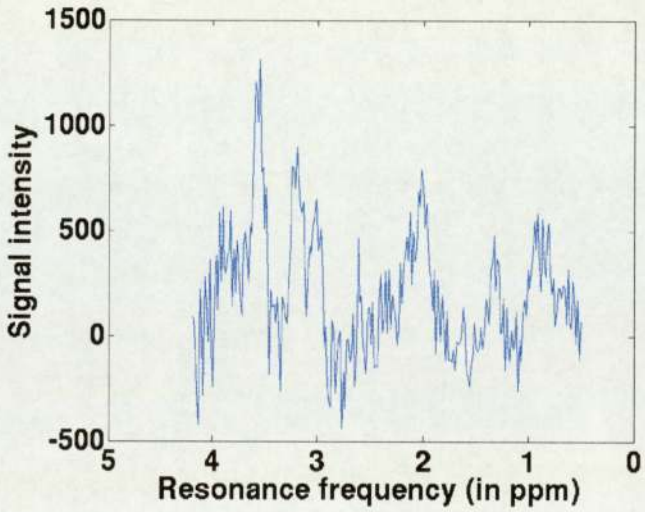


Figure 5.15: Voxel 140 of a patient with oligodendroglioma grade 2.

5.2.3 Conclusion

Though in most cases, MRI scans overlayed with NeuroScale distances give the same kind of information as the classifiers, in some cases as those shown above NeuroScale can reveal some anomalous or tumorous regions that would be overlooked by classifiers.
Furthermore the NeuroScale approach gives more stable results than the classification approach as the classifiers' predicted probabilities of being tumorous for each voxel vary greatly even if the classifiers have the same sensitivity, specificity, AUC and classification threshold.

Chapter 6

Conclusions

The predominant approach to a diagnosis problem such as brain tumour diagnosis is a classification approach. Several methods have been assessed in this task (Simonetti, 2004, Devos, 2005, Devos et al., 2005, Lukas, 2003). Applying visualisation techniques on the data before doing any classification can help find the most suitable method to discriminate the data.
But using visualisation, and more particularly topographic mapping, has yet other benefits when applied to the brain tumour diagnosis problem. Because it relies on the pairwise relationships between data and helps explore these relationships in a rather straightforward way (through the computation of distances between the data points of interest and a chosen reference point), it can uncover abnormalities of behaviour (in our case the chemical behaviour) in the data points considered with respect to a reference (that is deviations from a reference considered as normal), which would go unnoticed with a classification approach as shown in chapter 5 on page 79.
This property has been used in this thesis to detect anomalous areas outside the tumour region that would be classified as healthy but have a chemical behaviour suggestive of a tumorous process. To make it easy to spot these anomalous regions, the MRSI chemical information has been overlayed on MRI scans in a sort of visual "fusion" of both data as described in section 3.3 on page 47.
There are still a number of issues about the topographic mapping approach used in this thesis that have to be investigated.
The first one is the choice of the metric to use to generate the NeuroScale models and compute the overlayed distances between data points of interest and reference point. A corollary of this issue would be the investigation of the influence of using prior subjective class knowledge (α parameter) on the MRI scans overlayed with NeuroScale distances. Other issues related to these, and suggested by the glioma versus meningioma case, would be how to incorporate spatial information knowledge and how to perform normalization on the data in order to improve the quality of the overlayed scans.
Another important issue would be how to determine accurately the anomalous zones on the overlayed scans and especially their boundaries.
Finally, the work done on the healthy versus tumorous class (to help detect the existence of a

tumour) should be extended to all the other discrimination cases. This includes the delimitation of the region for which a grading of the tumour or a determination of the type of the tumour should be made and then using the same procedure as the one described for the healthy versus tumorous class on the data of the delimited region and using the discrimination case that has to be tested.

Appendices

Appendix A

Visualisation methods

A.1 NeuroScale

The principle of Neuroscale (Lowe and Tipping, 1997, Tipping and Lowe, 1997, Tipping, 1996) is to map p-dimensional data into a q-dimensional space, q being lower than p in general (q is taken equal to 2 in our study to be able to visualise the points corresponding to data on a 2- dimensional space), using a radial basis function neural network approach in an architecture known as Neuroscale. Neuroscale is a machine learning functional embodiment of a Sammon map, optimised to minimise a STRESS criterion function. The inputs are the relative dissimilarities between patterns in p, and the outputs are 'co-ordinates' in the q-dimensional space. The transformation is accomplished such that the overall relative distances between points in q-space reflect the relative dissimilarities specified in the p space.

Training is different to a traditional use of a neural network since there are no targets specified as such, but it is not completely unsupervised. The term *relative supervision* has been used for the training process. See Tipping and Lowe, 1997 for further details.

In addition, since a radial basis function network is used to effect the mapping, once optimised we then have a mechanism to project any new patient data through the architecture to inspect its visualisation position in the new topographic feature space.

A NeuroScale model may also include some knowledge about the classes to which the p-dimensional data points we want to map belong if it is available, by augmenting the distance function with some class knowledge.

A.1.1 Traditional NeuroScale

If d^*_{ij} is the 'distance'[1] between two data points of the dataset in the p-dimensional input space ($\mathbb{R}^p$) and d_{ij} is the distance between these same two points in the q-dimensional output space, ($\mathbb{R}^q$), then an ideal mapping of these data points would mean that

$$d^*_{ij} = d_{ij} \tag{A.1}$$

But given the reduction of dimensions to produce the mapping, this equality of distances is not always possible unless the original data genuinely was distributed on a 2-D manifold in the p-dimensional space. Therefore obtaining the best mapping possible for a given dataset is obtained by minimising the unnormalised STRESS (STandardised REsidual Sum of Squares) function

$$E = \sum_{i}^{N} \sum_{j}^{N} (d^*_{ij} - d_{ij})^2 \tag{A.2}$$

[1] 'distance' can be defined by any measure of dissimilarity, subject or objective, local or global, and need not even satisfy the properties of 'distance' such as symmetry, or positivity.

(Note that in traditional Sammon maps there is an additional normalising denominator employed which we ignore here). The 'distances' in the input and output space will be chosen to be either Euclidean or Minkowski in this work, though prior knowledge may dictate more appropriate choices. If the output of each point $\mathbf{x}_i$ and $\mathbf{x}_j$ is defined using an RBF thus using a parameterised non-linear function than to each input point $\mathbf{x}_i$ and $\mathbf{x}_j$ would correspond an output point

$$\mathbf{y}_i = f(\mathbf{x}_i, \mathbf{w}) \tag{A.3}$$

where $\mathbf{w}$ is a weight vector. Therefore the stress function to minimise would become :

$$E = \sum_i^N \sum_j^N (d_{ij}^* - ||f(\mathbf{x}_i, \mathbf{w}) - f(\mathbf{x}_j, \mathbf{w})||)^2 \tag{A.4}$$

This expression may then be differentiated with respect to the weight vectors in order to minimise the stress function E. A more appropriate optimisation strategy known as *shadow targets* was discussed in (Tipping and Lowe, 1997).

This basic approach can be augmented in several ways: for instance the incorporation of some prior class information knowledge by augmenting the data space dissimilarity measure with a class-dissimilarity measure. Let s_{ij} be the class-dissimilarity measure between two data points (such as the city-block distance between binary vectors representing different classes) in the input space. A modified STRESS function can be given by

$$E = \sum_i^N \sum_j^N (\delta_{ij}^* - ||f(\mathbf{x}_i, \mathbf{w}) - f(\mathbf{x}_j, \mathbf{w})||)^2 \tag{A.5}$$

where

$$\delta_{ij}^* = (1 - \alpha) d_{ij}^* + \alpha s_{ij} \tag{A.6}$$

α is a parameter in the interval $[0, 1]$ that controls the degree to which our prior knowledge of the classes influences the output representation. However in this thesis, in most cases, we limit the investigation to $\alpha = 0$, ie without using any class knowledge of the obtained data points. Hence the distribution of points in the visualisation is driven purely from the measured MRS sensor responses or from the measured MRI sensor responses or the combination of the measures from both sensors.

A.2 GTM

The goal of the GTM (Generative Topographic model) is to find a representation of the distribution $p(\mathbf{t})$ of the latent space of variables (of q dimensions) which generates the p-dimensional data we are trying to visualise, through a non-linear function plus noise. The distribution $p(\mathbf{t})$ of the latent space is chosen to be a constrained mixture of Gaussians whose parameters are determined using an EM algorithm. Magnification factors can be computed for this model and define the stretching and compression of the manifold representing the latent space, when embedded into the data space. These magnification factors can for example help uncover instances of multi-modality in the posterior distribution and further enhance the possible detection of data clusters. For more details see Bishop, Svensen and Williams, 1998, Svensen, 1998.

Appendix B

ROC curve analysis

B.1 A few preliminary definitions

If we define a two-class problem and are particularly interested in one of the two classes (for example the healthy class in the healthy versus tumorous problem), we would call true positives (respectively false positives) the data points truly belonging to this class and correctly identified as belonging to it by a classifier (respectively the data points not belonging to this class and falsely identified as belonging to it by a classifier). We can also define true negatives (respectively false negatives) as the data points not belonging to this class and correctly identified as not belonging to it by a classifier (respectively the data points belonging to this class and falsely identified as not belonging to it by a classifier).

We then define the sensitivity of a classifier as being the proportion of points of the class of interest correctly identified as belonging to the class of interest and the specificity of a classifier as being the proportion of points not belonging to the class of interest correctly identified as not belonging to the class of interest. That is more formally:

$$\text{sensitivity} = \frac{\text{true positives}}{\text{true positives} + \text{false negatives}}$$

$$\text{specificity} = \frac{\text{true negatives}}{\text{true negatives} + \text{false positives}}$$

B.2 ROC curves and AUC

The sensitivity and specificity of a classifier depend not only on the "quality" of the classifier but on the definition of the cutpoint between two classes. If we want either to minimize false negatives or false positive, we may wish to use different cutpoints.
The Receiver Operating Characteristic curve (or ROC curve) the plot of the true positive rate (sensitivity) against the false positive rate (1-specificity) for the different possible cutpoints between two classes. The area under the ROC curve (or AUC) is used as an indicator of the accuracy of a given classifier. An area of 1 represents a perfect classifier; an area of 0.5 represents a worthless classifier. The traditional academic point system can help determine the accuracy of a classifier:

- AUC between 0.90 and 1 : excellent classifier
- AUC between 0.80 and 0.90 : good classifier
- AUC between 0.70 and 0.80 : fair classifier

- AUC between 0.60 and 0.70 : poor classifier
- AUC between 0.50 and 0.60 : worthless classifier

In other words, an AUC of 1 represents perfect prediction, that is all positive cases sorted above all negative cases.An AUC of 0.5 is random prediction; there is no relationship between the predicted values and truth. An AUC below 0.5 indicates there is a relationship between predicted values and truth, but the model is backwards, that is, predicts smaller values for positive cases.

Appendix C

The use of hierarchical clustering in brain tumours diagnosis

The first method that can be tried in order to classify voxels in tumorous or healthy and then grade the tumorous voxels is hierarchical clustering. In the figure C on the next page, representing the results of hierarchical clustering for healthy versus tumorous, we can see that the only variables that may be informative are NAA and lactate.

A small group appears to have high values of lactate (yellowish shades in the column lactate). This might correspond to the group of high grade tumours. Another group appears to have slightly higher NAA values than the other groups (slightly brighter red shades), this should correspond to the healthy group but the difference in shades is too insignificant to draw any firm conclusion. All the groups not mentioned above cannot be identified by any relevant chemical information and we may only infer what they can be in contrast with the groups mentioned above.

Conclusion

Using hierarchical clustering on the data available to us is non informative. Other methods should be used to diagnose brain tumours correctly.

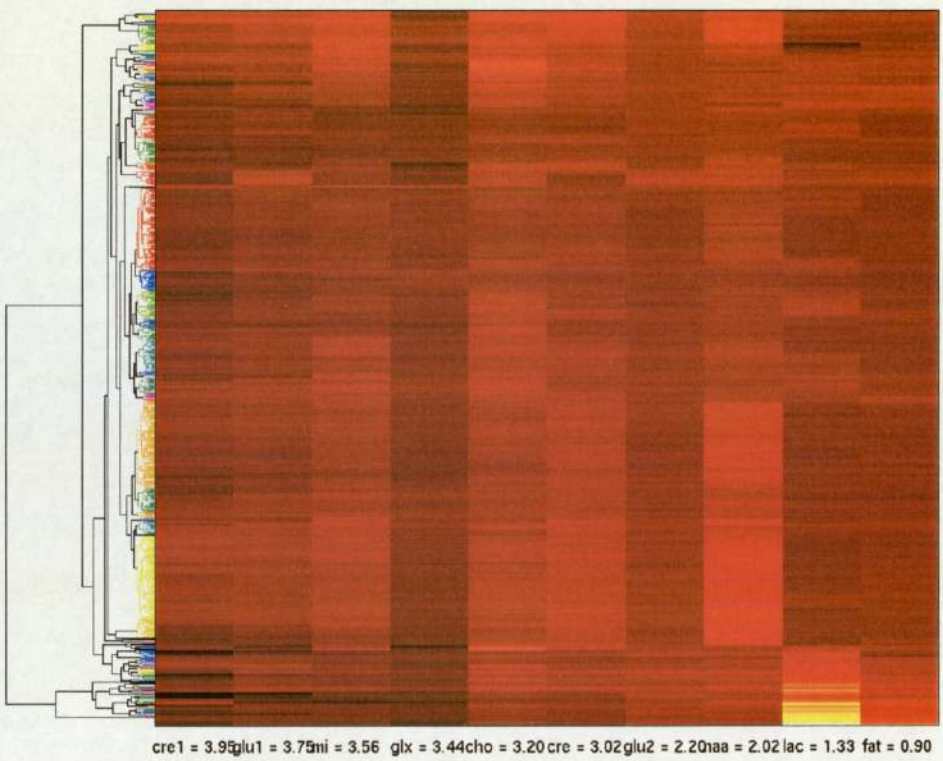


Figure C.1: Hierarchical clustering of healthy versus tumorous

Appendix D

Complete results of the classification experiments

Table D.1: Performance of classifiers using peak integration values data

	RBF with holdout	RBF with Leave-one-out	MLP with holdout	MLP with leave-one-out
Healthy versus tumorous	0.7445-0.9718	0.8683-0.9814	0.000000	0.0469-0.0878
Glioma versus meningioma	0.9461	0.9814	0.9833-0.9952	0
Low Grade Glioma versus High Grade Glioma	0.9279-0.9944	0.9586-0.9955	0.000000	0-0.0172
Low Grade Tumour versus High Grade Tumour	0.9567-0.9828	0.9673-0.9872	0.4	0.007-0.0104
Glioma Grade 2 versus Glioma Grade 3	0.6529-0.9278	0.8072-0.9233	0.000000	0.149-0.4298
Astrocytoma Grade 2 versus other grade 2	0.6355-0.7987	0.7547-0.822	0.000000	0-0.1812
Oligoastrocytoma Grade 2 versus other grade 2	0-0.7656	0-0.886	0.000000	0.1554-0.2214
Oligoastrocytoma Grade 3 versus other grade 3	0.4809-1	0.7204-1	0.000000	0-0.8547
Oligodendroglioma Grade 3 versus other grade 3	0.482-0.6601	0.605-0.9594	0.000000	0-0.1625

Table D.2: Performance of classifiers using MRI image intensities data

	RBF with holdout	RBF with Leave-one-out	MLP with holdout	MLP with leave-one-out
Healthy versus tumorous	0.7273-0.7756	0.7733-0.7929	0.000000	0.0625-0.1792
Glioma versus meningioma	0.8464-0.8562	0.864-0.8843	0.000000	0.2129-0.3797
Low Grade Glioma versus High Grade Glioma	0.7418-0.8739	0.8735-0.8983	0.000000	0.1346-0.2582
Low Grade Tumour versus High Grade Tumour	0.7736-0.8529	0.7449-0.8891	0.000000	0.1357-0.3054
Glioma Grade 2 versus Glioma Grade 3	0.5402-0.7824	0.5338-0.79	0.000000	0.4475-0.5929
Astrocytoma Grade 2 versus other grade 2	0.2315-0.604	0.3961-0.6733	0.000000	0.3256-0.4916
Oligoastrocytoma Grade 2 versus other grade 2	0.152-0.5351	0.1318-0.498	0.000000	0-0.1944
Oligoastrocytoma Grade 3 versus other grade 3	0.3229-0.6354	0.3793-0.7081	0.000000	0.4169-0.4791
Oligodendroglioma Grade 3 versus other grade 3	0.2614-0.4003	0.2406-0.5331	0.000000	0.3594-0.4319

Table D.3: Performance of classifiers using the combination of peak integration values data and MRI image intensities data

	RBF with holdout	RBF with Leave-one-out	MLP with holdout	MLP with leave-one-out
Healthy versus tumorous	0.6571-0.9969	0.968-0.9986	0.000000	0.1853-0.7488
Glioma versus meningioma	0.9726-0.9926	0.9722-0.9964	0.18	0.059-0.2018
Low Grade Glioma versus High Grade Glioma	0.9477-0.9984	0.95-0.9956	0.000000	0.0042-0.1863
Low Grade Tumour versus High Grade Tumour	0.9392-0.9925	0.9572-0.9948	0.000000	0.0107-0.2605
Glioma Grade 2 versus Glioma Grade 3	0.7512-0.8858	0.8667-0.9462	0.000000	0.2789-0.6086
Astrocytoma Grade 2 versus other grade 2	0.5285-0.8114	0.6355-0.805	0.000000	0.3432-0.4183
Oligoastrocytoma Grade 2 versus other grade 2	0.1896-0.8043	0.1298-0.8422	0.000000	0-0.1992
Oligoastrocytoma Grade 3 versus other grade 3	0.5486-0.7361	0.6502-1	0.000000	0-1
Oligodendroglioma Grade 3 versus other grade 3	0.4265-0.7157	0.6306-0.9988	0.000000	0-0.04

Appendix E

About metrics

The Minkowski distance of order p between two points i and j in a n-dimensional space is defined by:

$$d_{ij} = \left[\sum_{i=1}^{n} |x_{ik} - x_{jk}|^p \right]^{\frac{1}{p}} \tag{E.1}$$

k being the index of the coordinates.
When $p = 1$ then the distance measure is called the city block distance or Manhattan distance.For $p = 2$, the Minkowski metric gives the Euclidean distance.
The Minkowski distance does not account for different metrics of the individual coordinates. If the coordinates span different ranges, the results will be dominated by the coordinate with the largest range. Therefore, the data has to be scaled before calculating the distance.
Furthermore, the distances will be distorted if there is any correlations between variables (coordinates).
The Mahalanobis distance allows for correlation and different scalings of the variables and should be used whenever problems of scaling and correlation between the variables appear.

The Mahalanobis distance is related to the Euclidean distance, and gives the same results for uncorrelated and standardized data. It is defined by :

$$d_{ij} = \sqrt{(x_i - x_j)^T C^{-1} (x_i - x_j)} \tag{E.2}$$

Where C^{-1} is the inverse covariance matrix of the data.

Bibliography

Asavaphatiboon, S., Sinlapawongsaa, T., Laothamatasa, J., Dhanachaia, M., Theerapancharoenb, V. and Putthicharoenratc, S.: 2004, Comparison of mr spectroscopy and mr perfusion in benign and malignant infiltrative brain tumors, *in* K. D (ed.), *Radiosurgery*, Vol. 5, Basel, Karger, pp. 51–65.

Bishop, C., Svensen, M. and Williams, C.: 1998, The generative topographic mapping, *Neural Computation* **10**(1), 215–234.

Devos, A.: 2005, *Quantification and classification of Magnetic Resonance Spectroscopy data and applications to brain tumour recognition*, PhD thesis, Katholieke Universiteit Leuven, Leuven, Belgium.

Devos, A., Simonetti, A., van der Graaf, M., Lukas, L., Suykens, J., Vanhamme, L., Buydens, L., Heerschap, A. and van Huffel, S.: 2005, The use of multivariate mr imaging intensities versus metabolic data from mr spectroscopic imaging for brain tumour classification, *J. Magn. Reson.* **173 (2)**, 218–228.

http://www.medcyclopaedia.com/library/: n.d., GE Healthcare Limited. Medical Encyclopedia. **URL:** *http://www.medcyclopaedia.com/library/*

Kurhanewicz, J., Vigneron, D. B., and Nelson, S. J.: 2000, Three-dimensional magnetic resonance spectroscopic imaging of brain and prostate cancer, *Neoplasia* **2**, 166–189.

Lowe, D. and Tipping, M.: 1997, Neuroscale: Novel topographic feature extraction with radial basis function networks, *in* M. . Mozer, M. Jordan and T. Petsche (eds), *Advances in Neural Information Processing Systems 9*, Cambridge, Mass: MIT Press.

Lukas, L.: 2003, *Least squares support vector machines classification applied to brain tumour recognition using magnetic resonance spectroscopy*, PhD thesis, Katholieke Universiteit Leuven, Leuven, Belgium.

Preul, M., Caramanos, Z., Collins, D., Villemure, J., Leblanc, R., Olivier, A., Pokrupa, R. and Arnold, D.: 1996, Accurate, noninvasive diagnosis of human brain tumours by using proton magnetic resonance spectroscopy, *Nature Medicine* **2**, 323–325.

Saporta, G.: 1990, *Probabilités, analyse de données et statistique*, Technip.

Simonetti, A.: 2004, *Investigation of brain tumor classification and its reliability using chemometrics on MR spectroscopy and MR imaging data*, PhD thesis, University of Nijmegen, Nijmegen, Netherlands.

Svensen, M.: 1998, *GTM: The Generative Topographic Mapping*, PhD thesis, Aston University, Birmingham, UK.

Tipping, M.: 1996, *Topographic mappings and feed-forward neural networks*, PhD thesis, Aston University, Birmingham, UK.

Tipping, M. and Lowe, D.: 1997, Shadow targets: A novel algorithm for topographic projections by radial basis functions., *NeuroComputing Journal* **Special Issue on Radial Basis Functions**, 211–222.