UNIVERSIDADE DE LISBOA

FACULDADE DE CIÊNCIAS

DEPARTAMENTO DE FÍSICA

Automatic Reporting of TBI Lesion Location in CT based on Deep Learning and Atlas Mapping

Carolina Piçarra

Mestrado Integrado em Engenharia Biomédica e Biofísica
Perfil em Biofísica Médica e Fisiologia de Sistemas

Dissertação orientada por:
Dr. Ben Glocker
Dra. Raquel Conceição

2021

Acknowledgements

Abstract

The assessment of Computed Tomography (CT) scans for Traumatic Brain Injury (TBI) management remains a time consuming and challenging task for physicians. Computational methods for quantitative lesion segmentation and localisation may increase consistency in diagnosis and prognosis criteria. Our goal was to develop a registration-based tool to accurately localise several lesion classes (i.e., calculate the volume of lesion per brain region), as an extension of the Brain Lesion Analysis and Segmentation Tool for CT (BLAST-CT).

Lesions were located by projecting a Magnetic Resonance Imaging (MRI) labelled atlas from the Montreal Neurological Institute (MNI MRI atlas) to a lesion map in native space. We created a CT template to work as an intermediate step between the two imaging spaces, using 182 non-lesioned CT scans and an unbiased iterative registration approach. We then non-linearly registered the parcellated atlas to the CT template, subsequently registering (affine) the result to native space. From the final atlas alignment, it was possible to calculate the volume of each lesion class per brain region. This pipeline was validated on a multi-centre dataset (n=839 scans), and defined three methods to flag any scans that presented sub-optimal results. The first one was based on the similarity metric of the registration of every scan to the study-specific CT template, the second aimed to identify any scans with regions that were completely collapsed post registration, and the final one identified scans with a significant volume of intra-ventricular haemorrhage outside of the ventricles. Additionally, an assessment of lesion prevalence and of the false negative and false positive rates of the algorithm, per anatomical region, was conducted, along with a bias assessment of the BLAST-CT tool.

Our results show that the constructed pipeline is able to successfully localise TBI lesions across the whole brain, although without voxel-wise accuracy. We found the error rates calculated for each brain region to be inversely correlated with the lesion volume within that region. No considerable bias was identified on BLAST-CT, as all the significant correlation coefficients calculated between the Dice scores and clinical variables (i.e., age, Glasgow Coma Scale, Extended Glasgow Outcome Scale and Injury Severity Score) were below 0.2. Our results also suggest that the variation in DSC between male and female patients within a specific age range was caused by the discrepancy in lesion volume presented by the scans included in each sample.

Keywords: traumatic brain injury, computed tomography, lesion localisation, lesion segmentation

Resumo

Os traumatismos cranioencefálicos (TCEs) constituem uma das principais causas de morte a nível mundial e em todas as idades, com uma incidência de cerca de 69 milhões de casos e uma carga económica de 400 mil milhões de dólares americanos por ano. A tomografia axial computorizada (TAC) é a modalidade de imagem mais utilizada não só para a avaliação inicial das lesões, como também para gestão do tratamento a realizar. análise de TACs com este objectivo ainda é, na prática clínica corrente, uma tarefa morosa e desafiante para médicos e radiologistas.

Vários métodos computacionais têm sido desenvolvidos ao longo dos últimos anos para segmentação e localização de lesões, no sentido de aumentar a eficiência destes processos e a coerência dos critérios de diagnóstico e prognóstico. No entanto, os métodos desenvolvidos apresentam algumas limitações, incluindo apenas um tipo específico de lesão, ou indicando apenas em que lobo do cérebro a lesão se encontra. O intuito deste projecto é, desta forma, desenvolver uma ferramenta baseada no registo de imagens de TAC para localizar quantitativamente várias classes de lesões de TCEs (hemorragia extra-axial (HEA), hemorragia intra-parenquimal (HIP), hemorragia intra-ventricular (HIV) e edema) em 31 regiões cerebrais. A ferramenta proposta surge como uma extensão da *Brain Lesion Analysis and Segmentation Tool for CT* (BLAST-CT), uma ferramenta de segmentação automática de lesões de TCE em TAC, baseada numa rede neuronal convolucional.

A localização das lesões foi feita através da projecção de um atlas de ressonância magnética, no espaço do *Montreal Neurological Institute* (MNI), para um mapa de lesão, resultante da segmentação da mesma, que se encontra no espaço específico de cada sujeito. Tendo em conta a diferença entre os espaços em que as duas imagens se encontram, foi necessário criar um *template* de TAC, que funciona como um passo intermédio entre os dois espaços. Este *template* foi criado com 182 tomografias sem lesões detetáveis, e várias iterações de registo de imagem. Inicialmente, todas as imagens foram linearmente (*affine*) registadas para o espaço padrão, ponderado em T1, do MNI, sendo posteriormente alinhadas e calculada a sua média, de forma a criar uma primeira versão do *template*. De seguida, realizaram-se seis iterações de registo: três *affine* e três não lineares. Em cada iteração, todos os *scans* foram registados para o *template* criado na iteração anterior, sendo depois alinhados e a sua média calculada, formando um novo *template*, mais nítido do que o anterior. Em comparação com o *template* de TAC anteriormente disponível, o resultado deste processo é significativamente melhor, com contornos mais precisos e melhor alinhamento das estruturas anatómicas.

O atlas segmentado foi então registado, de forma não linear, para o espaço do *template* de TAC, sendo o resultado posteriormente registado (*affine*) para o espaço específico de cada sujeito. Como o atlas foi segmentado no mesmo espaço do mapa de lesão, foi possível sobrepor as duas imagens e calcular não só o volume de cada classe de lesão por região cerebral, como também os volumes totais do cérebro

e de cada região. Esta metodologia foi validada com um *dataset* constituído por 839 exames de TAC com lesões detetáveis, oriundos de 38 centros europeus. Para cada exame estava disponível a segmentação automática, feita pelo BLAST-CT, e uma segmentação de referência, conduzida por médicos especialistas. Este *dataset* será referido como *Dataset* 2. Todos os dados utilizados são provenientes do estudo Collaborative European Neuro Trauma Effectiveness Research in TBI (CENTER-TBI).

De forma a sinalizar os exames cuja localização poderá ter tido resultados abaixo do esperado, três métodos foram definidos. O primeiro baseou-se na métrica de similaridade, calculada no registo do atlas segmentado (no espaço do *template* de TAC) para o mapa da lesão, e que classifica a qualidade do mesmo. Foram sinalizados todos as imagens com uma métrica abaixo de 0,65, tendo sido este um limiar definido empiricamente. No segundo método foram utilizados os volumes de cada região anatómica, calculados para cada paciente, para construir um *box-plot* por região, com o objetivo de identificar os *outliers* de *box-plot*. Os exames que apresentassem mais de cinco regiões com volumes que constituíssem *outliers* foram inspecionados visualmente, sendo desta forma identificados os resultados sub-ótimos. Por fim, foram identificados quaisquer exames que apresentassem um volume significativo (>1mL) de hemorragia intra-ventricular fora dos ventrículos do cérebro. 96 exames foram identificados com o primeiro método, 12 com o segundo e 10 com o terceiro. 9 dos exames identificados com o segundo método foram igualmente identificados pelo primeiro, o que indica que o segundo método poderá ser redundante.

Adicionalmente, com o objectivo de analisar quais as regiões anatómicas mais afetadas por cada tipo de lesão, foram construídos mapas visuais de prevalência, utilizando os volumes da localização das lesões nos exames do *Dataset* 2, excluindo os exames previamente identificados como tendo resultados "sub-ótimos". Estes mapas destinaram-se também a apoiar a análise de erros de localização, descrita em seguida, dado que a gravidade de uma taxa de erro específica varia em função da prevalência da lesão numa dada região. Assim, para maximizar a sua precisão, todos os mapas de prevalência foram construídos utilizando os volumes por região obtidos com os mapas de lesão de referência (segmentados por médicos especialistas). Para criar cada mapa de prevalência (um por classe de lesão), foi inicializado um contador para cada região do atlas. Percorrendo todos os sujeitos, o contador é incrementado por um se o sujeito tiver um volume de lesão naquela região superior a um limiar definido. Foram aplicados dois limiares: 0,1mL e 1mL. O primeiro limiar foi estabelecido no sentido de excluir pequenas falhas no alinhamento do atlas, tendo em conta que com a metodologia apresentada não é possível atingir uma exatidão ao nível do voxel. Foi verificado que a prevalência de lesões de HEA, HIP e edema é significativamente maior na metade anterior do cérebro, enquanto as lesões de HIV são, por outro lado, mais prevalentes nos ventrículos. As lesões de HEA também apresentam, como seria esperado, maior prevalência em regiões contíguas com a fronteira cerebral.

Dois mapas de erro foram posteriormente construídos para cada tipo de lesão, com os valores médios das taxas de volumes de falsos negativos e falsos positivos, por região cerebral. Um limiar de 0,1mL foi aplicado a estes mapas, à semelhança do que foi feito aos mapas de prevalência acima mencionados. Assim, cada *scan* com um volume de referência - ou de forma equivalente, uma soma do volume verdadeiro positivo e falso negativo - inferior a 0,1mL, foi excluído do cálculo da média. Foi possível verificar que a taxa de volume falso negativo é, em geral, inversamente proporcional ao volume de lesão por região. Esta relação poderá dever-se quer a uma baixa performance do BLAST-CT, que se sabe ser influenciada pelo volume de lesão, quer à falta de precisão da ferramenta de localização. Na localização de lesões tão pequenas como 0,2mL numa região anatómica, um pequeno erro de registo

pode levar a uma localização incorreta. Durante toda as análises realizadas anteriormente, considerou-se que uma lesão estava presente numa determinada região quando ocupava um volume superior a 0,1mL. No entanto, os altos valores da taxa de volume falso negativo ($>0,75$), obtidos em regiões onde o volume mediano da lesão é baixo ($<0,7$mL), indicam que este limiar deverá ser aumentado.

Por fim, conduziu-se uma análise de correlação parcial entre a performance do BLAST-CT, medida pelo *dice score* de cada tipo de lesão, e várias variáveis independentes (idade, *Glasgow Coma Scale*, *Extended Glasgow Outcome Scale* e *Injury Severity Score*). O coeficiente de Pearson foi calculado para os testes em que a variável independente é a idade, uma vez que tanto esta variável como os *dice scores* são quantitativos e contínuos. Para todos os restantes testes, o coeficiente de correlação de *Spearman* foi utilizado, pois as restantes variáveis independentes podem ser consideradas ordinais e quantitativas. Em todos os testes estatísticos, entre cada par de variáveis, todas as restantes variáveis independentes foram incluídas como covariáveis, bem como o tempo desde a origem da lesão, o volume da lesão e o sexo biológico do sujeito. Não foi identificado qualquer viés na performance do BLAST-CT, uma vez que todos os coeficientes de correlação estatisticamente significativos, calculados entre os *dice scores* e as variáveis clínicas, estavam abaixo de 0,2. Os resultados também sugerem que a variação do *dice score* detetada entre pacientes do sexo masculino e feminino, dentro da mesma faixa etária, é causada não pela diferença de sexo biológico, mas sim pela discrepância no volume de lesão apresentado pelos exames incluídos em cada amostra.

Em suma, observou-se que a metodologia proposta é capaz de localizar lesões de TCE com sucesso, em toda a área cerebral, embora a sua precisão não chegue ao nível do voxel. O trabalho futuro poderá incluir uma análise da evolução e consistência da localização entre exames múltiplos, i.e., do mesmo sujeito mas em momentos diferentes, de forma a investigar as diferenças e os padrões espaciais de mudança para cada tipo de lesão.

Palavras-chave: lesão cerebral traumática, tomografia computorizada, localização de lesão, segmentação de lesão

Index

List of Figures

LIST OF FIGURES

List of Tables

List of Acronyms

BLAST-CT Brain Lesion Analysis and Segmentation Tool for Computed Tomography

CC Cross-correlation

CENTER-TBI Collaborative European Neuro Trauma Effectiveness Research in TBI

CNN Convolutional neural network

CSV Comma-separated values

CT Computed Tomography

DOF Degrees-of-freedom

DSC Dice similarity coefficient

EAH Extra-axial haemorrhage

FNR False negative rate

FPR False positive rate

GCS Glasgow Coma Scale

GD Gradient descent

GOSE Glasgow Outcome Scale–Extended

HEA Hemorragia extra-axial

HIP Hemorragia intraparenquimal

HIV Hemorragia intra-ventricular

ICH Intracranial haemorrhage

ISS Injury severity scale

IPH Intraparenchymal hemorrhage

IQR Interquartile range

IV Independent variable

IVH Intraventricular haemorrhage

MI   Mutual information

MNI   Montreal Neurological Institute

MRI   Magnetic Resonance Imaging

NCCT Non-contrast CT

NN   Nearest neighbour

ROI   Region of interest

SM   Similarity metric

TAC   Tomografia Axial Computorizada

TBI   Traumatic Brain Injury

TCE   Traumatismo Cranioencefálico

Chapter 1

Objectives

Traumatic Brain Injury (TBI) remains a severe under-recognised public health challenge. Currently, it is one of the leading causes of disability and death globally across all ages, with an estimated incidence of around 69 million cases [1] and an economic burden of 400 billion US dollars per year [2]. Traffic accidents and falls are the most significant contributors to TBI incidence in both high income and low to middle-income countries, being therefore expected to increase with demographic ageing [2, 3].

TBI is defined as a form of acquired brain injury caused by an external force. Despite this simplistic definition, it is a complex condition that can be branched into several sub-classes and results in various cognitive, emotional, psycho-social and behavioural outcomes [4].

Computed Tomography (CT) remains the gold standard imaging modality not only for initial assessment of TBI but also for treatment guidance due to its ease of access, short acquisition time and sensitivity for the detection of acute haemorrhagic lesions, which may require immediate and targeted intervention [5, 6]. The increased sensitivity of Magnetic Resonance Imaging (MRI) for the detection of non-hemorrhagic contusions, diffuse axonal injury, cytotoxic oedema, and microhaemorrhages also make it a promising modality for the identification of some TBI subtypes, which are often missed by CT. However, limitations such as access, potential instability, long acquisition time and sensitivity to motion limit its use in emergency settings [4], thus being usually reserved either for an evaluation of neurologic findings not explained by the CT scan, or as a research tool [6, 7].

The quantitative assessment of CT scans for TBI management is a time consuming, expensive and challenging task that requires specialised radiological expertise, which is often missing due to a lack of efficient and accessible concussion training for physicians [8]. Nonetheless, it is crucial to provide valuable information for patient management, prognosis, and pathology analysis/monitoring [9, 10].

Several retrospective studies have presented evidence of a significant rate (16.7% [8], 56.0% [11] and as high as 76.9% [12]) of mild TBI (mTBI) misdiagnosis in emergency departments (EDs). The inter-observer variability and absence of standardisation regarding both diagnosis and symptomatological analysis criteria - which may reinforce the effect of physicians' confirmation bias - are posed as impactful factors for this lack of accuracy [12]. Additionally, the identification of TBI patients at high risk for non-acute repercussions (e.g. higher healthcare usage, psychosocial consequences and neurodegeneration) versus TBI patients who are safe to discharge, also represents a challenge for clinicians.

1. OBJECTIVES

The use of automatic methods for medical imaging analysis (lesion segmentation and localisation) may increase consistency in diagnostic criteria by quantifying the extent of intracranial abnormalities, improving routine clinical practices and prognosis models, and significantly decreasing the radiologists' work time per exam. This is of particular need in low to middle-income countries, where the considerably higher TBI incidence is accompanied by a low availability of experts [4, 7].

The main goal of this project is thus to extend an automatic lesion detection and segmentation tool in order to report clinically-useful quantitative metrics about lesion load and location. An additional analysis of lesion prevalence and of the algorithm error rate in each brain region was conducted, along with a bias assessment of the original lesion segmentation tool.

This project was proposed by Dr Ben Glocker, from the Biomedical Image Analysis Group (BioMedIA) at Imperial College London, UK. From the indicated external institution, it counts with the supervision of Dr Stefan Winzeck (Post-doctoral Researcher) and Miguel Monteiro (Research Assistant/PhD Student). Lastly, Dr Raquel Conceição, from the Institute of Biophysics and Biomedical Engineering at Faculdade de Ciências da Universidade de Lisboa, is responsible for the internal supervision.

This report is divided into three main chapters. In chapter 2, a literature and state-of-the-art review is presented, with a section on TBI and its current classification practices. It also includes an overview of recent computational methods developed for the assessment of brain lesions in CT scans, as well as of image registration methods. The following chapter goes over the full lesion localisation pipeline, including the construction of a CT template. In chapter 4, the analysis of the performance of the localisation algorithm is presented, along with the bias assessment of the original lesion segmentation tool. Lastly, chapter 5 includes the conclusions of the dissertation.

Chapter 2

Literature review

2.1 Traumatic Brain Injury

The definition of TBI as a form of acquired brain injury caused by an external force leaves space for a wide diversity of clinical manifestations regarding the nature, intensity, location, direction and duration of the impact of the external force.

The first stage after trauma typically consists of a series of primary parenchymal damage followed by various systemic and local effects (e.g. hypoxia, hypotension, hypercarbia, axonal injury, brain swelling and compression due to increased intracranial pressure, and overall neurodegeneration), which combined lead to secondary brain damage [5]. The mortality of severe TBI is estimated at 30-40% [2], and the secondary effects can affect survivors several years after the trauma. Besides short term effects - as post traumatic amnesia and loss of conscience - TBI has been shown to be epidemiologically associated with an increase in the long-term mortality rate and in the risk for cognitive impairment and Alzheimer's disease [13, 14, 15, 16, 17], stroke [18, 19], epilepsy [20] and Parkinsonism [21, 22, 23, 24, 25], although the latter in reduced magnitude [26].

2.2 TBI classification and acute assessment

TBI is a complex condition due to its high patient and injury heterogeneity. The current diagnosis and lesion assessment practice presents several challenges related to the subtlety of neurological signs and symptoms in mTBI and the difficulty in making prognosis decisions on moderate and severe TBI. Robust evidence to support general guidelines and recommendations is lacking, leading to the need for local protocols and health professionals' expertise to fill these knowledge gaps, consequently decreasing consistency across centres [2, 4].

In the last two decades, TBI management guidelines have evolved towards a generalising approach, not considering the clinical variability between patients [4]. However, a significant part of current research reflects the concept of precision medicine, which has been shown to be effective in several medical fields, including neurocritical care [27]. Achieving a better and more personalised characterisation (e.g. identification of subtype, volumetric quantification and spatial localisation) of this condition at an acute state would theoretically enable a more accurate diagnosis and treatment management, improving clinical outcome [4, 7].

2. LITERATURE REVIEW

2.2.1 Current clinical practice

The clinical manifestations of TBI depend greatly on the origin of the external force. The first distinction to be commonly made is between closed and open head injuries, as the management principles significantly differ between the two classes. Furthermore, the initial assessment of severity - classifying TBI as mild, moderate or severe - is conventionally completed according to the individual's level of consciousness and/or duration of amnesia [28], resulting in the Glasgow Coma Scale (GCS) [29]. The GCS presents a rapid severity evaluation of brain injuries and is assessed by measuring spontaneous and stimulated verbal, motor and eye opening responses [30]. It may be employed at the scene of the injury as a tool to guide the triage process or at the ED. This classification tool is often considered mildly insufficient, as it does not take into account the mechanistic variability and different sub-types of TBI. Besides, its application has been shown to be confounded by the use of pre-hospital sedation, tracheal intubation and intoxication, and weakly correlated to the 12-month outcome [4, 31, 32].

Additionally, one of the most commonly used TBI-specific prognostic scores is the Glasgow Outcome Scale–Extended (GOSE) [33]. It is a scale of functional outcome based on a structured questionnaire, which rates patients state into one of eight categories: Dead, Vegetative State, Lower Severe disability, Upper Severe Disability, Lower Moderate Disability, Upper Moderate Disability, Lower Good Recovery or Upper Good Recovery. Although it is currently the gold-standard measure of global outcome after TBI, it presents several limitations: it includes only one scale item to assess mild/low-severity TBI and lacks quantitative and more precise variables to assess the multi-dimensional nature of TBI outcome [33].

2.2.1.1 Neuroimaging and classification tools

As mentioned above, neuroimaging methods such as CT and MRI are regularly used as tools to anatomically examine primary and secondary injuries, often in conjunction with GCS scores and other clinical information. In emergency settings, the goal of imaging is to identify, as soon as possible, lesions that need immediate surgical intervention or may benefit from early medical therapy or close supervision. Additionally, it is also useful for prognosis determination, to tailor rehabilitative therapy or assist the discharge planning [34, 35].

Despite its additional radiation exposure risk, for decades there has been a robust consensus that non-contrast CT (NCCT) is the first imaging scan performed in emergency settings for acute mild to severe TBI assessment [35, 36]. NCCT presents the main diagnostic advantage of being sensitive and specific for the presence of intracranial haemorrhage (ICH) (Figure 2.1 a.). This detection is crucial in early management as thrombolytic agents cannot be used on patients with ICH. NCCT is also sensitive to other injuries relevant for TBI assessment, as extra-axial fluid collections, skull fractures (Figure 2.1 b.), radiopaque foreign bodies, cerebral oedema (Figure 2.1 c.), swelling and signs of herniation [34, 35]. However, it does present low soft tissue contrast and consequently limited sensitivity for cerebral ischemia identification. In clinical practice, NCCT is widely used in EDs because of its low acquisition time and widespread availability. Furthermore, the screening of patients for ferromagnetic substances is not needed [35] and it is applicable to unconscious or intubated patients.

a.	b.	c.
Large acute left parieto-occipital parenchymal haemorrhage.	Left parietal bone fracture (arrowheads) with marked overlying soft tissue contusion.	Low attenuation is seen (arrows) representing interstitial oedema from transependymal flow of CSF.
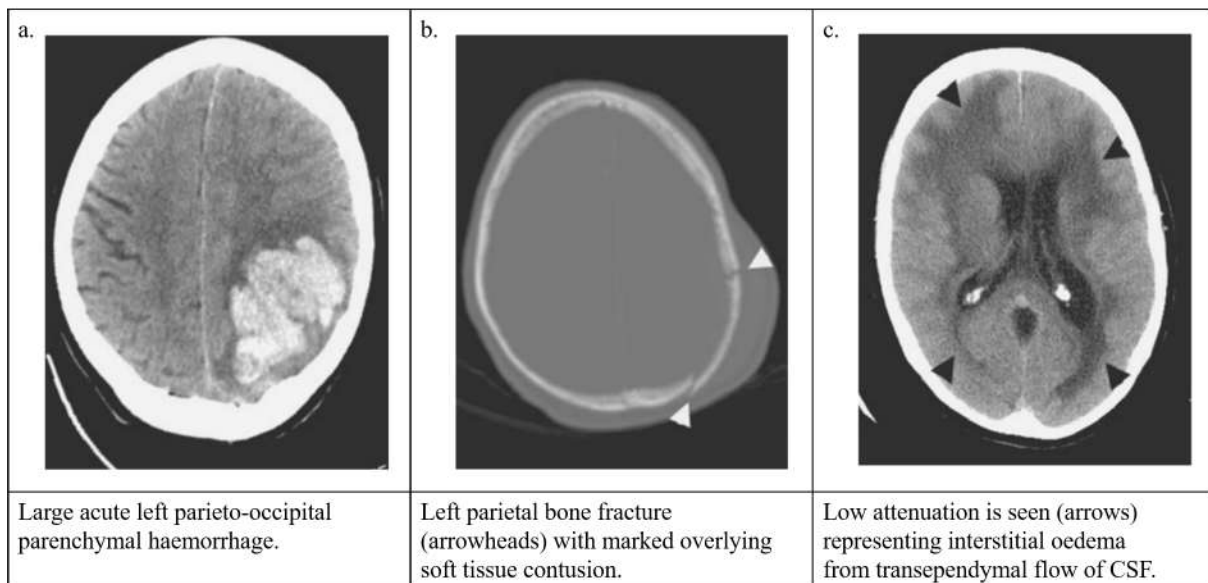
Figure 2.1: Common TBI injuries on CT. Images from [37]

2.2.1.2 Scoring systems based on CT findings

Regarding classification, several methods allow the employment of early CT findings to determine prognosis among patients with moderate to severe TBI. The most used method is the Marshall classification scheme [38], an anatomically oriented tool that stratifies patients into six different groups solely based on the type and severity of several CT scan abnormalities [39]. It focuses on the status of the basal cisterns, degree of midline shift, presence of a mass lesion, and criteria for surgical intervention [40]. Despite being frequently used, research has shown mixed results regarding its predictive power [40, 41, 42]. Besides, it is susceptible to inter-observer variability and limited regarding patients with more than one injury. The Rotterdam CT classification [39] consists of a refinement of this scheme, based on the combination of NCCT findings, such as the presence of haemorrhages and degree of mid-line shift [35], and has been shown to have significant prognostic value [43, 44]. Another alternative severity scoring system is the Abbreviated Injury Scale (AIS), which categorises each injury by body region (intra or extracranial) according to its relative severity on a 6 point scale (1: minor and 6: maximal) and results on the Injury Severity Score (ISS) [45]. The latter is usually a retrospective score which has been shown to be significantly but weakly correlated with functional outcome [32].

2.2.1.3 Sub-types of TBI

Besides being classified by severity, the assessment of CT scans can also result in a categorisation of TBI according to the level and type of brain damage caused. The primary effects of TBI can be focal or diffuse, according to its origin mechanism. Regarding focal injuries, there are several possible occurrences:

- **Skull fractures** indicate a strong impact and can be classified as linear, depressed, or base of skull fractures. Linear fractures often do not require treatment, contrary to depressed brain fractures, which are usually interventioned to reconstitute the skull's structure and avoid further brain damage, depending on the location of the lesion. A skull fracture significantly increases the risk of having a subarachnoid, subdural or epidural haemorrhage, and of yielding widespread infection

2. LITERATURE REVIEW

if open. Its location should be reported per fracture as: Right, Left, Bilateral (i.e., single fracture extending across midline), midline, frontal, parietal, temporal, occipital, skull base [28, 43].

- **Epidural Haematoma** is a collection of blood between the skull and dura. A skull fracture typically accompanies it. Although it appears hyperdense on CT, hypodense areas may appear, representing unclotted blood [43].

- **Extra-axial Haematoma (subarachnoid, subdural or epidural)** is a collection of blood between the brain surface and the skull, which cannot be specifically classified or located.

- **Subdural Haematoma** is a collection of blood outside the brain, in the subdural space (between the arachnoid and the dura layer), usually uniform and hyperdense on CT. Similarly to an extra-axial haematoma, hypodense regions may appear if the lesion contains unclotted blood, CSF admixture or active extravasation, which usually happens over time. It can result in increased intracranial pressure if untreated, leading to loss of consciousness [28, 43].

- **Subarachnoid Haematoma** is defined as an agglomeration of blood between the brain surface and the arachnoid matter, which usually follows the contour of the sulci and cisterns. It is usually hyperdense in CT but sometimes not visible [43].

- **Midline Shift** is the displacement of the septum pellucidum caused by a space-occupying lesion or consequent brain swelling. It is usually measured at the foramen of Monro, or where it presents the highest intensity value [43].

- A **Cortical Contusion** consists of a focal area of brain parenchymal disruption, typically affecting brain regions of skull contact, where irregular bony protuberances are present. Unlike intracerebral haematomas, contusions usually do not present clearly visible and uniform haemorrhages, showing both haemorrhagic and non-haemorrhagic tissue [28, 43].

- **Intracerebral/intraparenchymal haemorrhages (IPH)** are space-occupying collections of blood inside the brain's parenchyma, resulting from vascular damage (caused by, for example, brain laceration or diffuse axonal injury). IPH lesions are surrounded by hypodense tissue, which may be oedema or a clot retraction. Haemorrhages can also be intraventricular if they are located within the ventricles [28, 43].

Additionally, lesions can also be diffuse:

- **Diffuse axonal injury** consists of several widespread small lesions (hemorrhagic or non-hemorrhagic) in white-matter (WM) tissue.

Finally, there are some relevant secondary effects of TBI worth mentioning:

- **Oedema** is considered a secondary effect of TBI. It is defined as a water accumulation in the intracellular and/or extracellular spaces, and it is always hypodense in CT. It is one of the leading causes of increased intracranial pressure and can be divided into four classes: cytotoxic, vasogenic, interstitial and osmotic. Cytotoxic oedema is caused by a disruption in cellular metabolism, resulting in an accumulation of sodium and consequently of water. In CT, the abnormality is usually located within WM. Contrarily, in vasogenic oedema the blood-brain barrier is disrupted, and the excess fluid is typically seen in WM. In interstitial oedema, fluid is present within the extracellular space of periventricular WM. Finally, osmotic oedema is caused by a slightly higher plasma osmolality

in relation to brain tissue. Independently of the type, oedema is usually found around the lesion (perilesional oedema). [28, 43]

- **Hypoxic-Ischemic Injury/Ischemia** includes a variety of findings in tissue that maintains a deficit between substrate demand and delivery, which may be caused by an infarction, arterial occlusion or a global insult that leads to hypoxia or hypotension. The location of this effect may have an important pathophysiological connotation, as it corresponds to a specific vascular area [28, 43].

In addition to the tools and qualitative classes mentioned in this section, a quantitative analysis of pathology in head CT should add relevant insights to the functional outcome of patients. However, this requires accurate manual lesion segmentation and determination of affected regions, which is, as mentioned, a deeply tedious, time-consuming and operator-dependent task. Tools that provide qualitative information, or at best approximate calculations of lesions volume and prevalence, are thus predominant in current clinical practice. The development of automatic quantitative CT lesion assessment methods has therefore become a major focus in TBI research over the last two decades [42].

2.3 Computational methods for lesion CT assessment and prognosis

2.3.1 State-of-the-art review

Recent research has carved the way towards an increasing acknowledgement that an accurate characterisation of lesions and outcome prediction may only be possible through a multidimensional and individualised approach, including multiple domains such as patient's genetic, blood, cerebrospinal fluid biomarkers along with clinical and various neuroimaging data [4, 7, 46]. Following this line of research, several computational methods have been developed to automatically quantify and characterise different lesion types, with the ultimate goal of developing a more consistent and reliable patient-centred pipeline for the initial assessment of acute TBI CT scans and prognosis models [7, 42]. Within this category, several tasks/goals can be discriminated: lesion segmentation, volumetric quantification, differential qualitative classification of lesions and lesion location (i.e. identification of affected anatomical regions).

Regarding intracranial lesions detection and segmentation on CT scans, early studies have explored the fact that contusions and blood have intensity values in the same range of the Hounsfield scale by applying intensity thresholding and clustering-based segmentation methods [47, 48, 49, 50, 51, 52, 53]. Yuh et al. [54, 55] developed an algorithm for the detection of five features: presence or absence of i) subdural or epidural hematoma, ii) subarachnoid haemorrhage, iii) intraparenchymal hematoma, iv) clinically significant midline shift (>5 mm), and v) normal, partly effaced or completely effaced basal cisterns. The detection of these features was, once again, based on CT density thresholds, spatial filtering, and cluster analysis. Overall, these methods may perform well on standard images but struggle with situations such as the overlap of the haemorrhage region with other brain tissue, or poorly defined lesion edges [56]. Contemporaneously, several studies relied on techniques as active contouring and level sets, which frequently require manual input (initialised seed) [48, 57, 58, 59], introducing some degree of inter-observer variability. Most of the referred studies either did not differentiate between different types of haemorrhages or only addressed a specific class of lesions. Furthermore, most had considerably small validation samples, which might raise concerns about the robustness of the algorithms and clinical application feasibility. Additionally, the desired accurate volumetric analysis of lesions is only possible with

2. LITERATURE REVIEW

voxel-wise (versus image or patch-level) approaches [7], which might be too computationally expensive for such methods.

On a different perspective, Koikkalainen et al. [60] attempted to improve outcome prediction by extracting quantitative features - as total blood volume, midline shift and ventricles asymmetry - from CT scans. This type of analysis helps improve prognosis. However, it does not allow for a quantitative characterisation of individual lesions, which is expected to lead to a better understanding of lesion progression and more clinically relevant prognostic schemes [7].

Over the last few years, advances in deep learning-based methods have demonstrated significant potential to extract clinically relevant information from medical images [61, 62] in several medical fields [63, 64, 65, 66, 67]. For segmentation tasks, deep convolutional neural networks (CNNs) present the advantage of being able to learn highly discriminative features, such as complex non-linear mappings between the input image and the segmentation map. Several recent studies employed deep learning methods (i.e. different architectures of CNNs) for detection [68] or segmentation [69, 70, 71, 72] of undifferentiated haemorrhagic lesions. Li et al. [71] proposed a 2D U-net-based deep learning framework to automatically detect and segment undifferentiated haemorrhage strokes in CT scans. The authors also assessed lesion location at an image level. However, the evaluation of this location was done by overlapping the lesion map with the ground truth map and performing a binary analysis in which any overlap value higher than zero led to the conclusion that the task had been successful, which might overestimate its real performance. Nonetheless, satisfactory results were obtained regarding the segmentation of larger lesions. Despite its limited value towards a precise quantification of lesions or for outcome prediction, this type of binary analysis might be valuable for the selection of patients that need immediate intervention [7].

Conversely, some studies have conducted a more specific analysis by developing algorithms for detection [73] and segmentation [61, 74] of differentiated haemorrhage types at image level, all with considerably satisfactory samples sizes. However, none of these studies used TBI data or presented quantitative values for lesion volume or location. Remedios et al. [75] presented a framework through which the same model - for segmentation and volume calculation of undifferentiated haematoma lesions in TBI CT scans - could be trained on multiple sites without the data being transferred between them. However, this study focused on comparing the results obtained with the training set from each site. Additionally, Yao et al. [76] and Farzaneh et al. [77] developed algorithms for automatic segmentation and volume quantification of undifferentiated haematomas and subdural haematomas, respectively. Both studies focused specifically on TBI. Yao et al. [76] also applied the volume and shape characteristics extracted to build a random forest model of 6-month-mortality. The predictive power of the extracted features was analysed using 10-fold cross-validation on a clinical trial dataset of 828 patients.

Recently, Monteiro et al. [7] developed a CNN-based TBI specific tool for voxel-wise detection and segmentation of four lesion types: intraparenchymal haemorrhage, extra-axial haemorrhage, intraventricular haemorrhage, and perilesional oedema. This tool will be addressed in greater detail in the next section. To the best of our knowledge, no other TBI focused articles have quantitatively analysed oedema lesions.

Considerable research has been completed towards finding the significant relation between head

lesion location and the functional outcome of patients [78, 79, 80, 81, 82, 83, 84, 85, 86, 87]. Isokuortti et al. [83] examined the distribution of subdural haematomas, subarachnoid haemorrhages and contusions in a representative sample (n=3023) of TBI patients' CT scans. However, this analysis was non-automatic and only qualitative, classifying a lesion as frontal, parietal, temporal or occipital. To the best of our knowledge, most of the studies addressing this topic performed qualitative assessments similar to [78, 80], and most of them were focused on stroke [84, 85, 87]. Chastain et al. [82] and Ernst et al. [85] conducted a more specific analysis by using atlas registration to calculate the overlap of the lesion volume with each brain region. Nonetheless, in [82], the results had low precision, as each lesion was simply assigned to the region containing the majority of its volume. Ideally, the output of a lesion location analysis would be a quantitative value such as volume, specific for every anatomical region, i.e. the volume of every region affected by the lesion.

2.3.2 Brain Lesion Analysis and Segmentation Tool (BLAST-CT)

The Brain Lesion Analysis and Segmentation Tool for Computed Tomography (BLAST-CT) consists of an algorithm based on deep CNNs for multiclass, voxel-wise segmentation, volumetric quantification and classification of TBI lesion types in CT scans [7]. Monteiro et al. [7] employed a previously developed three-dimensional CNN with three parallel pathways (DeepMedic [9, 42]), which processes the input at different resolutions. In order to validate an initial CNN, 98 CT scans from 27 patients were used. This CNN was then employed to automatically segment a second dataset, of 839 scans from 38 different centres. The resulting segmentation maps were refined by clinicians. A subgroup of this second dataset (184 scans) was subsequently used as a training set in the final CNN, and the remaining 655 scans formed the testing set to evaluate the model performance. All the data mentioned above were from the Collaborative European Neuro Trauma Effectiveness Research in TBI study (CENTER-TBI) [88]. Finally, an independent dataset (500 scans from the CQ500 dataset) was employed for external validation of the CNN.

Unlike previous studies on the same topic, Monteiro et al. [7] were able to accurately segment and determine the volume of four different classes of lesions observed in CT images: IPH, extra-axial haemorrhage (EAH), which includes subdural haematoma, extradural haematoma, and traumatic subarachnoid haemorrhage, intraventricular haemorrhage (IVH), and perilesional oedema. The use of voxel-wise labels also allows for the localisation of lesions, for which the authors proposed a pipeline. However, questions remain regarding the optimisation and clinical assessment of the lesion location task. Hence, our primary goal is to extend BLAST-CT in order to obtain a quantitative analysis of the affected anatomical regions (i.e., how much volume of each lesion class is present in each of the regions defined by the selected atlas). With this objective in mind, several image registration techniques will be used throughout this project, as will be described in Chapter 3.

2.4   Image registration

Image registration is defined as the process of spatially aligning two or more images with the same content taken at different times, from different viewpoints, or by different sensors, to the same coordinate space [89]. Thus, the general goal is to find the optimal transformation that aligns the structures of interest in the images to be registered [90]. Over the last 20 years, image registration has become

2. LITERATURE REVIEW

a significantly important tool in several clinical areas, such as radiotherapy - for diagnosis and tumour staging, for example - and neurosurgery - where multimodal registration is part of the pre-surgical patient management to optimise surgical planning [91].

The registration process can be classified according to different criteria:

- **Dimensionality (spatial or spatiotemporal 2D-to-2D, 2D-to-3D, 3D-to-3D)** - Based on the spatial dimension of the input data;

- **Source of the used image features** - Registration methods require anatomical information, such as apparent prominent landmarks (surfaces, curves or edges), binary divided structures or voxel image intensities. Depending on which type of features are used, methods can be referred to as landmark-/geometrical-, segmentation- and voxel-/intensity-based methods, respectively [91, 92, 93]. The present introduction will be focused on intensity-based methods as those were the ones used throughout the project.

- **Nature of the transformation (e.g. rigid, affine, projective or deformable methods)** - A rigid transformation only allows the sensed image to be translated and rotated with respect to the reference image, being therefore limited to 6 degrees-of-freedom (DOF) for the registration of 3D images. This transformation is commonly used as pre-registration, before a more complex transformation is applied. Alternatively, affine transformations also allow for scaling and shearing of the input images, being defined by 12 parameters ($[W]$) (3 translation, 3 rotation, 3 scaling and 3 shearing parameters). The transformation T is given by T(X) = WX + S, where S is the translation vector:

$$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} w_{11} & w_{12} & w_{13} & w_{14} \\ w_{21} & w_{22} & w_{23} & w_{24} \\ w_{31} & w_{32} & w_{33} & w_{34} \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} + \begin{bmatrix} s_1 \\ s_2 \\ s_3 \\ 0 \end{bmatrix} \tag{2.1}$$

Finally, deformable transformations can be useful when the two images present considerable local differences. In order to achieve an accurate result in such scenarios, it is necessary to establish a dense voxel-wise non-linear spatial correspondence, allowing for a much higher number of DOF. The goal is to maximise the similarity between the fixed and moved image - similarly to the remaining transformations introduced - while maintaining a smooth displacement field, and it can be defined as:

$$\phi^* = arg_\phi min f_{sim}(F, M(\phi)) + f_{reg}(\phi) \tag{2.2}$$

where F and M denote the fixed and moving images, respectively, ϕ^* designates the optimal displacement field (ϕ), f_{sim} represents the dissimilarity function and f_{reg} the smoothness regularisation function [94]. There are several types of deformable registration methods frequently applied to medical images, such as free-form deformations with b-splines [95], the optical flow-based Demons algorithm [96] and diffeomorphic registration methods as the symmetric image normalisation method (SyN) [97]. The first two examples given parametrise the optimisation problem with displacement fields, smoothed by an energy function or a Gaussian filter. However, the true inverse transformation of a displacement field is not guaranteed to exist. It is therefore important that the transformations are diffeomorphisms (i.e. invertible and differentiable map with a differentiable inverse), to guarantee smooth and one-to-one mapping and consequently preserve the
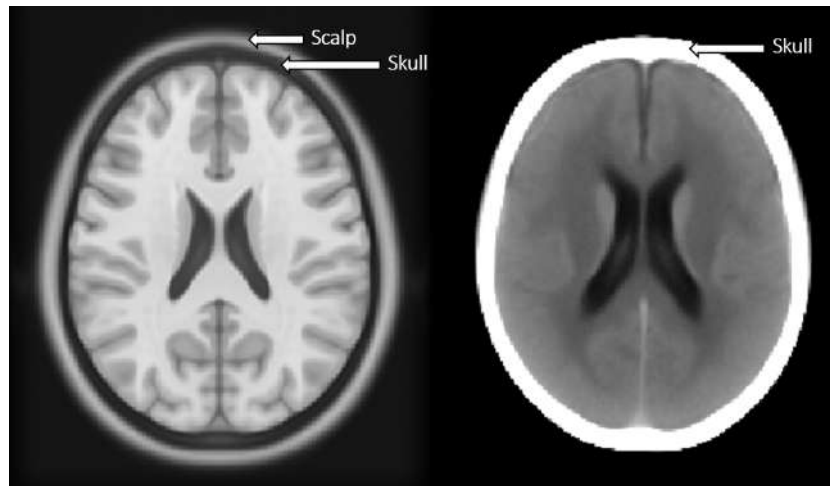
Figure 2.2: Structural differences between brain edges in MRI and CT. While on CT the higher intensity external layer corresponds to the skull, on MR it is comprised of mostly scalp tissue.

topology of the structures [94, 90]. The free-form and flow-based registration methodologies can also be diffeomorphic if a penalty term is added to the similarity measure or adequate constraints are applied in order to avoid undesirable deformations [90]. Klein et al. evaluated 14 different nonlinear deformation algorithms using 8 different error measures [98]. In this study, SyN was consistently shown to be one of the methods with higher accuracy across all subjects and label sets.

- **Degree of interaction (interactive, semi-automatic or automatic)** - In an automatic method, the user should only supply input data (e.g. image data or information about the image acquisition). In semi-automatic methods, the user can either initialise or steer the algorithm by rejecting or accepting the suggested registration hypothesis. Finally, in an interactive process, the user does the registration himself, assisted by software. [99]

- **Number of input images (pairwise, n=2 images, or groupwise, n>2 images)** [91].

- **Modalities involved (mono-modality, multi-modality, modality to model, patient to modality)** [91]. Multi-modal registration presents significant challenges due to the disparity between ventricles size and skull intensities on CT (bright) and MRI (dark) scans, as shown in Figure 2.2.

- **Optimisation procedure (parameters computed or searched for)** - The goal of image registration is finding a geometric transformation that brings one image into the best possible spatial correspondence with another image or physical space by optimising a registration criterion. The parameters that describe a geometric transformation can be computed directly or searched for. Direct computation of transformation parameters is only possible when the correspondences between the images' features are known, as in extrinsic and landmark-based methods. In intrinsic registrations, such point pairs are not available and must therefore be searched for. This search should be done iteratively by minimising the distance between corresponding feature sets (i.e. maximising similarity measures and minimising cost functions). [93]

- **Input data characteristics (intra-subject, inter-subject or image-to-atlas)** [91].

- **Anatomical structure involved (e.g. brain, heart, breast)** [91].

2. LITERATURE REVIEW

There are several common identifiable elements on intensity-based registration methodologies: the geometric transformation (described above), similarity metric (SM), optimiser, interpolator and, if possible, an initial/pre-registration transformation. Overall, the goal is to iteratively search for the parameters of a geometric transformation that optimises the similarity between the moving and fixed images. The way this search is conducted is defined by the optimiser. The role of the interpolator is, on the other hand, to resample the voxel intensity into the reference coordinate space [90].

The most frequently employed SMs are based on intensity differences and information theory. The Sum of Squared Differences (SSD) is computed assuming that the corresponding voxels/structures in both images should have the same intensities. The goal is therefore for it to be minimised. Conversely, Cross-Correlation (CC) is a metric based on the assumption that the corresponding intensities are linearly correlated. CC is defined as the scalar product of the two images (A and B), each interpreted as a long vector of intensities with N voxel locations x, and should be maximised:

$$CC = \frac{1}{N}\sum_x A(x)B(x) \tag{2.3}$$

Both SMs mentioned are mostly appropriate for mono-modal registration, as they are highly dependent on the intensity distributions of the images. Mutual information (MI) and its derived metrics (e.g. Mattes MI), which are based on information theory, have become highly relevant over the last two decades [100, 101, 102]. MI is a measure of statistical dependency, being particularly suited for multi-modal registration methods:

$$MI = H(A) + H(B) - H(A,B) \tag{2.4}$$

where H(Image) is the Shannon-Wiener entropy of the analysed image:

$$H(A) = -\sum_c p_A(c)log(p_A(c)) \tag{2.5}$$

with $p_A(c)$ as the probability that a voxel in image A has intensity c. $H(A,B)$ is the joint entropy, defined as:

$$H(A,B) = -\sum_c \sum_d p_{A,B}(c,d)log(p_{A,B}(c,d)) \tag{2.6}$$

where $p_{A,B}(c,d)$ is the probability that a voxel in the joint image (A,B) has intensity (c,d). MI should be maximised (i.e. entropy of each image should be maximal, as an image with zero entropy would have no value for registration, but the joint entropy should be minimal) [103]. It should be noted that the described SMs only take into account the relation between voxels, not considering possible relevant structural information.

As mentioned, the goal of the optimiser is to find the transformation parameters that result in the optimal SM value. There are several commonly used optimisation algorithms, such as the gradient descent and its variations, the downhill simplex method or Powell's method. The Powell's direction set method finds the minimum of an N-dimensional function by repeatedly minimising the function in one direction along a set of N different directions until the function stops decreasing. Each direction is initialised with its basis vector in parameter space, but after an iteration has been completed, that is, after the function has been minimised in every direction once, the overall distance moved in parameter space is taken as a new direction. Given that medical images usually do not have the same resolution in every direction,

the order in which each parameter is optimised affects optimisation performance and consequently, registration robustness. One disadvantage of this method is that if there is a high difference between the function's second derivative across directions, many iterations through all basis vectors will be needed to arrive to the minimum [104, 105].

Alternatively, the downhill complex begins by constructing a non-degenerate simplex (i.e. a geometrical figure with N+1 vertices encompassing a finite N-dimensional volume inside) from N+1 points. The iterative deformation and contraction of this simplex (with volume conservation) leads to the translation of its vertices to the function's minimum. The optimisation usually terminates when the magnitude of the distance moved in one step or the corresponding decrease in function value is smaller than a previously set threshold. [104, 105].

Contrarily to the two previous algorithms, gradient descent (GD) requires derivative calculation. This method moves in the direction of the negative gradient vector of the cost/SM function, with a certain step size:

$$\mu_{k+1} = \mu_k - a_k * \frac{\partial C(\mu_k)}{\partial \mu_k} \tag{2.7}$$

where $\frac{\partial C(\mu_k)}{\partial \mu_k}$ represents the derivative of the cost function in the position μ_k, and a_k is the learning rate at position k. The latter can be defined in several ways. When defined by the user, a_k can either be a constant or a decaying function of k: $a_k = \frac{a}{(k+A)^\alpha}$, with all parameters set by the user under the following constraints: $a > 0$, $A \geq 1$ and $0 \leq \alpha \leq 1$. Alternatively, a line search can be used, which attempts, in each iteration, to minimise the cost function along the gradient direction:

$$a_k = argmin_a C\left(\mu_k + a\frac{\partial C(\mu_k)}{\partial \mu_k}\right) \tag{2.8}$$

This method requires a significant increase of cost function evaluations and consequently of computational expense. Thus, an inexact version is frequently employed, where the searched learning rate only yields a sufficient reduction of C [104, 106, 95]. Adding the line search to this method has been shown to improve the convergence rate, at the cost of an increase in computational time [106]. Another variation of GD - with faster convergence - is the conjugate gradient method, which uses the previous direction to define a new conjugate direction of search [104, 107].

A multi-resolution pyramidal strategy is often used to implement the iterative optimisation algorithm. The parameters (i.e. number of pyramid levels, number of iterations in each level, smoothing sigmas and shrink factors) should be defined based on the resolution of the input data and the structure within the image relative to this resolution [108]. This strategy follows a coarse-to-fine approach, using the resolution pyramid to down-sample the reference and moving images and subsequently registering them from the lower (i.e. using only a fraction of the image voxels) to the higher resolution images. In each level, the transformation calculated in the previous step is used as the new initial registration [90]. This approach presents the main advantage of being computationally efficient, since most iterations are completed at the coarser resolution levels. It is also less likely to fall into a local extreme as the initial search is done on a coarse grid [109].

When a point from the input image is mapped to the target space, it is generally assigned to a point which is not a grid position. It is therefore necessary to interpolate that voxel intensity as a weighted

2. LITERATURE REVIEW

sum of neighbouring voxel intensities. This interpolation might influence registration accuracy and efficiency, and there are several ways to conduct it, such as with a nearest neighbour (NN), linear or a more complex windowed sinc scheme. NN interpolates the intensity at a non-integer voxel position by assigning the closest grid position voxel a weight of 1, while all other voxels receive no weight. Alternatively, a linear interpolator, as the name indicates, linearly interpolates the intensity of the non-grid voxel position, which is sufficient in most practical cases. The windowed sinc scheme presents the advantage of having minimum aliasing artefacts, despite being more computationally expensive. In this method, the interpolation kernel along each spatial direction is defined as a sinc function $\left(\frac{sin(\pi t)}{\pi t}\right)$ multiplied by a window function, such as the Hamming, Welch or Cosine windows [110, 111]. In order to increase computational efficiency, a simple interpolator is usually used for the optimisation step (e.g. linear or NN), followed by the employment of a more complex scheme such as the windowed sinc function to resample the input image [90].

Choosing the right pre-registration transformation is crucial for the success of a registration framework, as it decreases the probability of divergence or convergence to a local minimum [90]. The most common and straightforward methods for registration initialisation are using the identity transform or setting the centre of rotation and translation of the transform. The latter can be completed by setting the geometrical centre of the moving image as the initial centre of rotation, and the vector between the centres of the two images as the initial translation. This approach is based on the assumption that the regions of interest are centred on their respective images. Alternatively, the centre of mass (calculated using the moments of grey level values) of the moving image can be set as the centre of rotation, with the vector between the centres of mass of the two images being set as the initial translation [112]. This method assumes that the moments of the two elements are similar, which is usually not valid for multi-modal registration. Aligning the centres of the images is important due to the non-linear nature of rotation. The greater the distance between the two anatomical objects, the more sensitive our computations are to small rotation angles, which is disadvantageous for numerical stability. The ideal centre of rotation should minimise the distance to the farthest point of our region of interest (ROI) [113].

Chapter 3

CT template construction and integration

3.1 Concept

Lesions were located by projecting an MRI labelled atlas, generated in MNI space, to the prediction map of BLAST-CT (in subject-specific space). Given the disparity of imaging spaces between the elements to be registered and the aforementioned challenges that multi-modal registration presents, there is the need to use a CT template, which works as an intermediate step between native and MNI space. As there was no appropriate pre-existent CT template available, this had to be constructed.

Figure 3.1 summarises the lesion localisation process: initially, a pre-localisation step has to be completed, where the parcellated atlas is registered (affine + deformable) to the CT template/study-specific space (0). The following steps need to be completed with each scan to be localised. Native CT scans are affinely registered to the CT template (1), and the inverse of this affine transformation is employed to register the parcellated atlas (in study-specific space) to native space (2). This final registration allows for the overlap of the atlas and each BLAST-CT prediction map, from which the volume of lesion per atlas region, the full volume of each atlas region and whole brain are retrieved (3). Computing an affine transformation - versus a deformable transformation - presents the relevant advantage of being much more time efficient, taking only up to a minute. Despite the potential benefit in the accuracy of the atlas alignment when applying a deformable transformation, this process is likely to take more than 10 minutes per image, depending on the parameters used, which would not be feasible in a research context where it is necessary to process a large number of scans.

The full methodology and results of this pipeline will be described in detail in the present chapter (Section 3.3.2 - Construction of the study-specific CT template; 3.3.3 - How the affine + deformable transform of step 0 was constructed; 3.3.4 - How the rigid + affine transform of step 1 was constructed and the volumes of step 3 extracted).

3.2 Methods

3.2.1 Setting and datasets

The data used for this project were collected as part of the Collaborative European Neuro Trauma Effectiveness Research in TBI study (CENTER-TBI, NCT02210221). Patients were recruited at 60 different European centres between Dec 9, 2014, and Dec 17, 2017. Since imaging was conducted as part

3. CT TEMPLATE CONSTRUCTION AND INTEGRATION
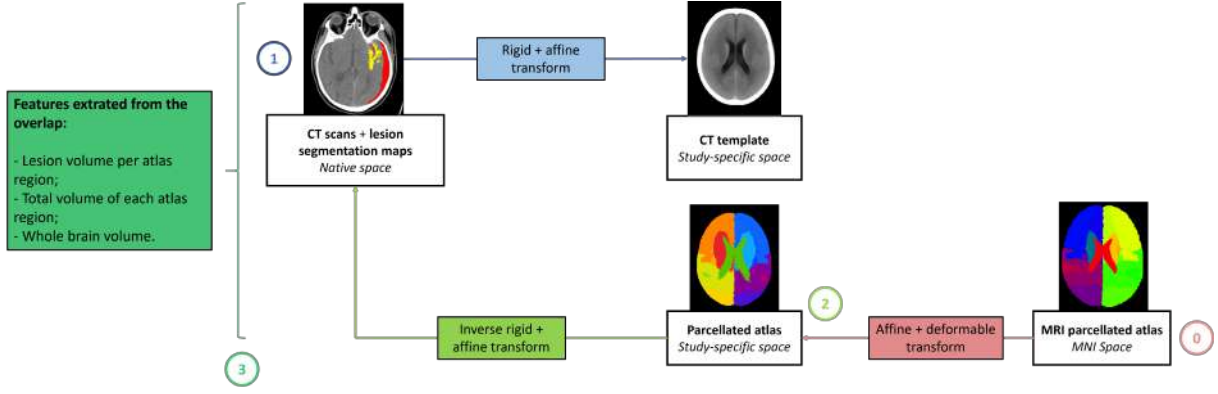
Figure 3.1: Flowchart of the full lesion localisation method. 0 - Parcellated atlas in MNI space is projected to the study-specific/CT template space; 1- Every native CT scan is registered to the CT template; 2- The inverse of the transformation calculated in step 1 is used to map the parcellated atlas in study-specific space to native space. 3- Relevant volumes are withdrawn from the overlap between the parcellated atlas and each lesion map.

of standard clinical practice, acquisition parameters were not standardised across sites [7].

The data were divided into two datasets. Dataset 1 consists of 189 TBI patients without abnormal findings and one scan per subject. Dataset 2 comprises 839 scans, acquired from 512 TBI patients with abnormal findings. Every scan of Dataset 2 was automatically segmented using a model based on manual segmentation by experts. These parcellations were further corrected by trained personnel, and 45 scans were double-checked in order to assess reproducibility and inter-rater variability [7]. Dataset 2 scans were also automatically segmented by BLAST-CT. These segmentations will be further referenced as "reference segmentation" and "prediction map from BLAST-CT", respectively. Four classes were identified in each parcellation: intraparenchymal haemorrhages, which also includes small petechial haemorrhages; extra-axial haemorrhages, which includes subdural haematomas, extradural haematomas, and traumatic subarachnoid haemorrhages; perilesional oedemas; and intraventricular haemorrhages.

3.2.2 CT template construction

Monteiro et al. [7] created a CT template in the context of the development of BLAST-CT. However, it included only 20 subjects from Dataset 1, with one subject selected as target and a four-step iterative registration process: one rigid, one affine, one deformable iteration, and a final affine alignment with the standard MNI MRI template. With the goal of improving its precision and anatomical accuracy, a new template one was constructed using all the scans on Dataset 1 (n=189) and an unbiased iterative registration approach, implemented with the *antsRegistration* function [108]. Every native scan was initially registered (affine) to the standard T1-weighted MNI template. The aligned scans were then averaged to create an initial CT template. Subsequently, six more registration iterations were conducted, three affine and three deformable. In each iteration, every native CT scan was registered to the preceding CT template (e.g. on the 2nd iteration, the scans were registered (affine) to the 1st/initial CT template, and on the 5th iteration, the native scans were registered (deformable) to the 4th template) and then averaged, creating a new sharper CT template to which the native scans were aligned in the following iteration.

Table 3.1 presents the registration parameters used for both the affine and deformable registration iterations. After each iteration, the MI SM was calculated between the registered scan and the template to which the images were registered, in order to identify scans with particularly low values. Scans with an

18

Table 3.1: Parameters employed when applying the *antsRegistration* function for the affine and non-linear registration of native CT scans to every target used during the CT template construction. SyN: symmetric image normalisation method.

Parameters	Affine registration	Deformable registration (SyN)
Dimensionality	3	3
Similarity metric	Mutual information Metric weight = 1 Nº of bins = 32 Sampling strategy: Regular Sampling percentage = 0.25	Cross correlation Metric weight = 1 Radius = 4 Sampling strategy = Regular Sampling percentage = 0.25
Interpolation	Welch windowed sinc	Welch windowed sinc
Optimiser	Gradient descent Gradient step = 0.2 —– —–	Gradient descent Gradient step = 0.1 Update field variance in voxel space = 3 Total field variance in voxel space = 0
Histogram matching	Yes	Yes
Convergence	Iterations per level: 200 x 100 x 50 Threshold = 10^{-6} Window size = 10	Iterations per level: 100 x 70 x 50 x 20 Threshold = 10^{-6} Window size = 10
Smoothing	4 mm x 2 mm x 1 mm	4 mm x 2 mm x 1 mm x 0 mm
Shrink factors	4 x 2 x 1	8 x 4 x 2 x 1

SM lower than 0.6[1] were visually inspected, resulting in the exclusion of seven scans due to consistently failed registrations, being cut-off or having motion artefacts.

3.2.3   CT template registration to MNI space

As shown in the schematic in Figure 3.1, the registration of the CT template to the MNI space is necessary in order to map the parcellated atlas to the study-specific space and facilitate the registration, and consequently the alignment, of native scans to the atlas.

The parcellated MNI atlas was previously constructed using 652 MR T1-weighted scans from the Cam-CAN study [114], parcellated via MALP-EM [115]. Patient-specific age-unbiased segmentations of cortical regions and ventricles were obtained through the projection of individual region atlases to the standard MNI MRI template. Distance maps were then used in order to map each WM voxel to its closest region, followed by the fusion of regions based on prior anatomical knowledge (e.g. medial frontal cortex was assigned to the medial frontal lobe). This resulted in a coarser subdivision of the brain into 31 different regions (Figure 3.2). A list of all the parcellated regions is available in the left column of Table B.1, shown in Appendix B.

The registration of the study-specific CT template to the standard MNI MRI template was completed with *antsRegistration*, whereby specific parameters (i.e. SM, sampling rate, number of iterations, shrink factors and smoothing sigmas) were empirically set. During the first phase of iterative parameter

---

[1]Threshold empirically defined.

3. CT TEMPLATE CONSTRUCTION AND INTEGRATION
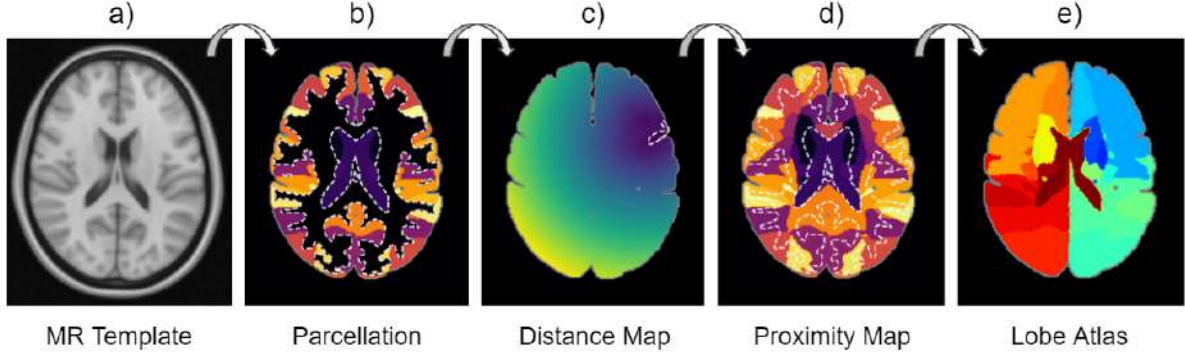
Figure 3.2: Method for the construction of the parcellated MNI atlas. Segmentations of cortical regions and ventricles were obtained through the projection of patient-specific region atlases to the standard MNI MRI template (b)). Distance maps (c)) were then used to map each WM voxel to its closest region (d)), followed by the fusion of regions based on prior anatomical knowledge (e)).

search, three combinations of SMs were tested for the affine and deformable registration: MI or Mattes MI (Mattes) as the SM from both the affine and deformable registration, and MI for affine and CC for deformable registration. All remaining parameters were kept constant and are displayed in the Appendix (Table A.1). After empirically choosing the most appropriate SM, all remaining parameters were iteratively optimised (Table A.2).

A final stage of parameter optimisation was then conducted with the goal of improving the skull alignment, in which Simple ITK *sitk.ImageRegistrationMethod* was used to construct the affine transformation, and *antsRegistration* was kept as the preferred tool for the deformable registration. To initialise the affine transformation, the *CenteredTransformInitializer* function was used to align the geometric centres of the two images. Mattes MI was set as the SM and the sampling strategy as regular, with a percentage of 25%. A gradient based optimiser was selected (Gradient Descent Line Search) with a learning rate of 1, 100 iterations, a convergence minimum value of 10^{-5} and window size of 5. Missing values were interpolated linearly. A three level multi-resolution pyramid was applied, setting the corresponding shrink factors and smoothing sigmas to [4, 2, 1] and [4, 2, 1] mm, respectively. By using this affine transformation as initialisation, a SyN transform was constructed with a gradient descent optimiser (learning rate= 0.1). The convolution kernel for both the update and total field variance are defined as 3 and 0 (specified in terms of the radius in voxel space), respectively. CC was set as the similarity metric with a metric weight of 1 and a radius of 5. No sampling strategy was defined. A four level multi-resolution registration approach was used, with 400, 200, 100 and 50 iterations for each level. The convergence threshold was set as 10^{-6} and window size as 5. The shrink factors and sigmas of gaussian smoothing for each level were defined as [6, 4, 2, 1] and [4, 2, 1, 0] mm, respectively.

Throughout the following sections, the result of this registration will be referenced as CT-MNI template.

3.2.4 CT template integration

To overlap the parcellated atlas (MNI space) regions with the lesion segmentation maps (subject-specific space), a transformation between the two imaging spaces had to be calculated. The previously constructed CT-MNI template was used as target instead of the MRI parcellated atlas, so the full regis-

tration pipeline could be mono-modal. Throughout the following methodology, the scans of Dataset 1 were used for initial testing due to the absence of lesions.

3.2.4.1   Registration of native CT scans to MNI space

As this transformation is meant to be applied to every image to be localised, it should be as computationally efficient as possible. Thus, direct affine registration between each native CT scan of Dataset 1 and the CT-MNI template was initially attempted. However, as the results were unsatisfactory, it was concluded that a two-step registration pipeline was necessary. The employed pipeline included rigid + affine registration of native CT scans to the study-specific CT template (fully completed with SimpleITK *ImageRegistrationMethod*) and the application of the previously calculated affine + deformable transformation between study-specific CT template and MNI space (Section 3.2.3).

The rigid Euler 3D transform was initialised by aligning the two images' geometric centres using the *CenteredTransformInitializer* function. Mattes MI was set as the SM and the sampling strategy as regular, with a percentage of 20%. A gradient based optimiser was selected (Gradient Descent Line Search) with a learning rate of 0.1, 200 iterations, a convergence minimum value of 10^{-6} and window size of 5. Missing values were interpolated linearly. A three level multi-resolution/pyramidal registration approach was used, setting the corresponding shrink factors and smoothing sigmas to [4, 2, 1] and [4, 2, 1] mm, respectively. The following affine registration was constructed using the same parameters, except for the initialisation (i.e. SM, sampling, optimiser, interpolator and multi-resolution level parameters). The previously found rigid transformation was set as the moving initial transform, while an empty affine transformation (created using Simple ITK's *AffineTransform(dimension)*) was set as the initial transform to be modified during the optimisation process.

This registration pipeline was both applied to Dataset 1 and 2. As the registration may not be robust for some cases with deep abnormalities, each native scan of Dataset 2 was registered four independent times. Only the most successful result, identified by the highest SM value, was kept. The SM values were used as a coarse guide for the analysis and visual inspection of the registration results, leading to the establishment of an SM threshold, below which the result was considered to be sub-optimal. The extraction of the volumes of lesion per atlas region was completed for all scans of Dataset 2, including the ones identified as sub-optimal. However, the SM value of each scan was kept as a feature to be considered for further quality control and statistical analysis. Additionally, the correlation between the SM values and lesion volume was calculated using the Pearson correlation coefficient.

3.2.4.2   Lesion localisation

The presented method allows for lesion localisation to be completed in both native and MNI space. In this study, the localisation and further analysis was fully done in native space. With computational efficiency in mind, the MRI atlas and its corresponding mask were projected *a priori* to the study-specific CT template using the transformation described in Section 3.2.3 (Step 0 of Figure 3.1). The rigid + affine transformation (Section 3.2.4.1) was then the only one applied on the actual localisation code (Step 3 of Figure 3.1), registering the atlas in study-specific space to native space. Once we have both the atlas and prediction map from BLAST-CT in native space, a mask of each atlas region is created and individually overlapped with the prediction map to extract the volume of lesion (differentiated by lesion class) present

21

3. CT TEMPLATE CONSTRUCTION AND INTEGRATION
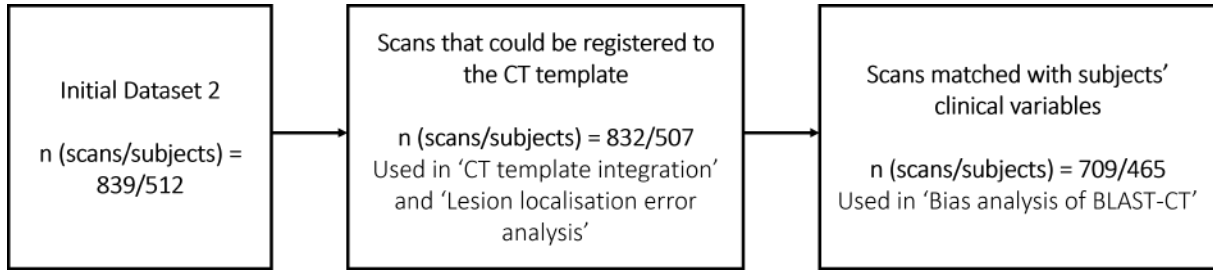
Figure 3.3: Schematic summary of the scans used in Dataset 2 throughout the several phases of the project.

in each region.

The output of this overlap was stored in a comma-separated values (CSV) file with the volume of lesion in each atlas region, separated by lesion class. Additionally, the full volume of each atlas region, as well as of the whole brain, were also calculated. All volumes were presented in millilitres. This lesion localisation methodology was fully implemented using the Python programming language [116] and the following packages: os (https://docs.python.org/3/library/os.html; last acessed: 1/08/2021), SimpleITK (simpleitk.org; last accessed: 31/08/2021), numpy (numpy.org; last accessed: 31/08/2021), pandas (pandas.pydata.org; last accessed: 31/08/2021) and operator (https://docs.python.org/3/library/operator.html; last acessed: 1/08/2021). When passing a CT scan through the localisation code, the user is able to save several extra registration parameters (e.g. number of iterations of the optimiser and the final SM values for both rigid and affine registration, the CT scan resampled to the CT template space and the analogous transformation), as well as to set the number of runs of registration to be completed, and if the localisation is done in native or atlas space.

3.3 Results

3.3.1 Experimental setup and datasets

Dataset 2 initially consisted of 839 scans, acquired from 512 patients in 38 different centres. However, only 832 of these were used in the development of the localisation algorithm. Furthermore, it was only possible to match 709 scans (465 patients) to their corresponding clinical information (Figure 3.3), which constituted the sample used for the bias assessment of the performance of BLAST-CT, presented in the following Section 4.1.2. Table 3.2 shows the descriptive statistics for clinical and TBI outcome variables of both datasets used, presented separately for women and men. The information displayed refers to every scan in Dataset 1 and 709 of the total 839 scans in Dataset 2. Count and percentages are presented for categorical variables while the median and corresponding range are used to describe continuous variables.

3.3.2 CT template construction

Figure 3.4 shows the final study-specific CT template alongside the previously template used in the context of the development of the initial version of BLAST-CT [7]. It is possible to observe that the new CT template is significantly sharper, with more precise outlines and overall better aligned anatomical regions. The ventricles, lateral sulci and interhemispheric fissure are more defined, as well as the cerebellum and brain stem, which are now well separated from the remaining cerebral tissue. We can also

Table 3.2: Cohort demographic and clinical information for both datasets used. The information displayed regarding Dataset 2 was obtained from only 709 scans out of the initial 839 scans. Percentages regarding Injury Severity Score do not add up to 100 because of rounding.

	Dataset 1		Dataset 2	
	Female	**Male**	**Female**	**Male**
$n_{subjects}/n_{scans}$	101/101	88/88	156/241	309/468
Age (years)	55 (14-95)	55 (6-89)	62 (10-89)	57 (6-89)
Glasgow Coma Scale	—-	—-		
13-15 (Mild TBI)			105 (67%)	185 (60%)
9-12 (Moderate TBI)			19 (12%)	30 (10%)
9 (Severe TBI)			26 (17%)	81 (26%)
Missing values			6 (4%)	13 (4%)
Extended Glasgow Outcome Scale at 6 months	—-	—-		
8 (Upper good recovery)			34 (22%)	82 (27%)
7 (Lower good recovery)			33 (21%)	41 (13%)
6 (Upper moderate disability)			19 (12%)	43 (14%)
5 (Lower moderate disability)			17 (11%)	35 (11%)
4 (Upper severe disability)			8 (5%)	14 (5%)
3 (Lower severe disability)			9 (6%)	41 (13%)
2 (Vegetative state)			0 (0%)	0 (0%)
1 (Death)			22 (14%)	34 (11%)
Missing values			14 (9%)	19 (6%)
Injury Severity Score	—-	—-		
1-8 (Minor injury)			35 (22%)	41 (13%)
9-15 (Moderate injury)			38 (24%)	58 (19%)
16-24 (Serious injury)			28 (18%)	54 (17%)
25-49 (Severe injury)			42 (27%)	118 (38%)
50-74 (Critical injury)			9 (6%)	23 (7%)
75 (Maximum injury)			2 (1%)	5 (2%)
Missing values			2 (1%)	10 (3%)
Lesion volume (mL)	—-	—-	1.95 (0.00-235.64)	3.84 (0.00-208.00)
Time since injury (h)	—-	—-	5.03 (0.27-608.78)	5.21 (0.33-1600.91)

3. CT TEMPLATE CONSTRUCTION AND INTEGRATION
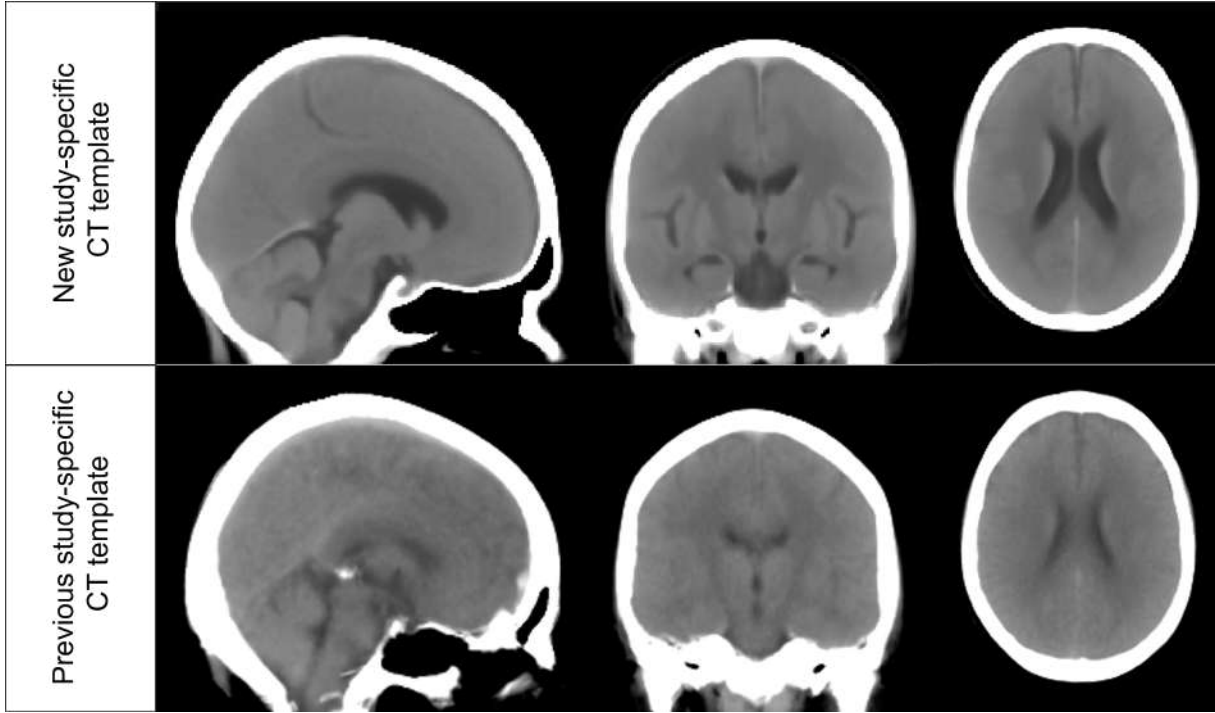
Figure 3.4: Final study-specific CT template, resulting from the 7th registration iteration, and previously created CT template, constructed in the context of the development of the initial version of BLAST-CT.

notice the absence of abnormal hyperintense structures in the frontal part of the brain and in the base of the brain stem.

3.3.3 CT template registration to MNI space

Figure 3.5 presents the different results of the registration of the study-specific CT template to the standard MNI MRI template, obtained with different SMs: a) MI or b) Mattes MI as the SM of both the affine and deformable registration, or with c) MI for affine and CC for deformable registration. All remaining parameters were kept constant and are presented in Table A.1 (Appendix A).

There seems to be a trade-off between achieving a satisfactory alignment of soft-tissue (particularly, the ventricles) or of the skull. Although none of the results presented in Figure 3.5 were highly acceptable, we can conclude that while using MI or Mattes may improve the skull alignment for both the affine and deformable registrations, it is only possible to achieve slightly better soft tissue alignment with CC. Given that the matching of the parcellated atlas to the lesion segmentation maps is based on affine registration (Step 3 of Figure 3.1), it is nonetheless necessary to expand the regions of the atlas. Thus, soft tissue alignment was prioritised and MI + CC were used as the SMs in further optimisation.

After an additional parameter search, the optimal set of parameters achieved combined a 5-levels pyramidal scheme with an increased number of iterations for deformable registration and heavy initial down-sampling. This approach allows, in theory, for better local alignment while avoiding excessive deformation of the images. However, although an acceptable soft tissue alignment was achieved, it was at the cost of a significant deformation of the top part of the skull, as can be observed in Figure 3.6. The final parameters used in this optimisation phase are shown in Table A.2 (Appendix A).
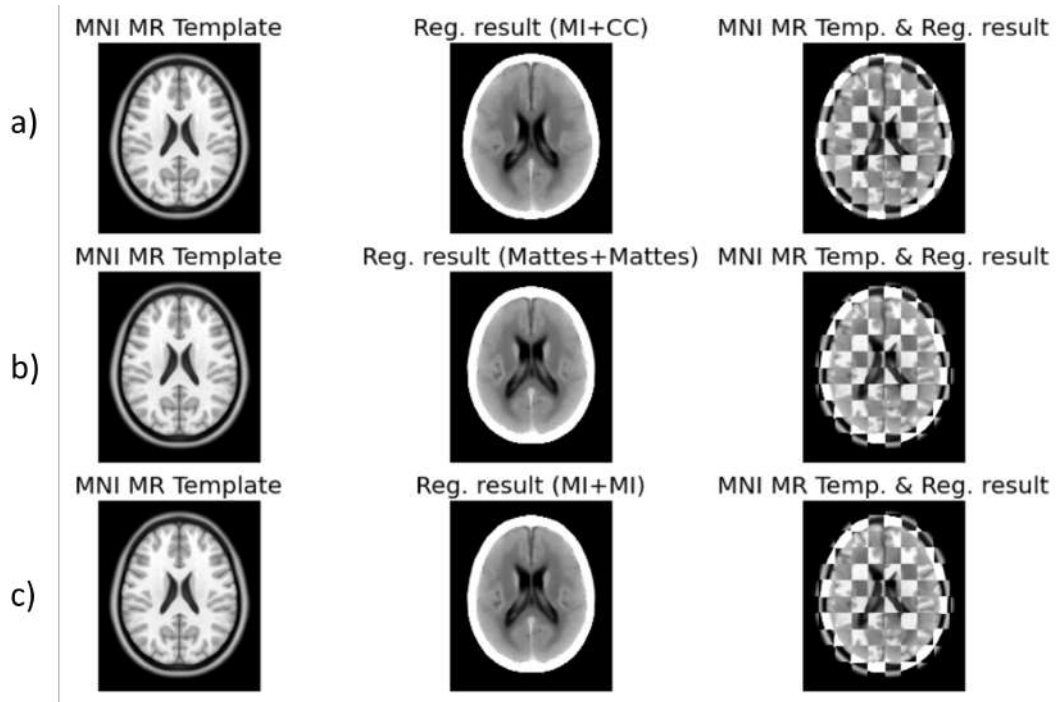
Figure 3.5: Qualitative results of the registration of the study-specific CT template to the standard MNI MRI template. All registration parameters were kept constant except for the SM (Table A.1, in Appendix A). a) MI and CC used as SMs for the affine and deformable registrations, respectively; b) Mattes MI used as the SM for both registration processes; c) MI used as the SM for both registration processes.
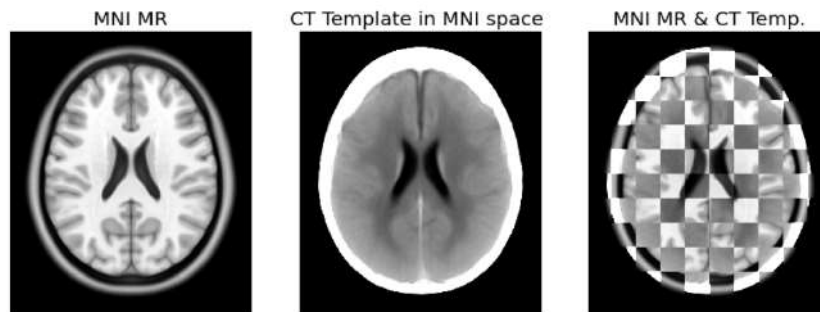
Figure 3.6: Qualitative registration results. Registration of the study-specific CT template to the MRI MNI atlas, after the second phase of iterative parameters optimisation. All registration parameters displayed in Table A.2 (Appendix A

3. CT TEMPLATE CONSTRUCTION AND INTEGRATION
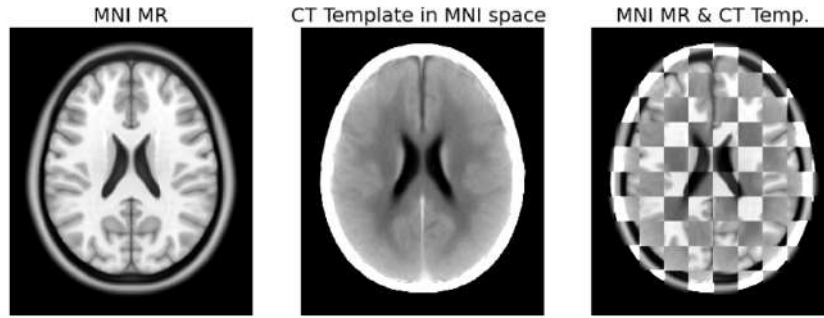
Figure 3.7: Final qualitative registration result, further referred as CT-MNI template. Registration of the study-specific CT template to the MRI MNI atlas.

With the goal of further improving the skull alignment, a new registration pipeline was employed, combining Simple ITK *sitk.ImageRegistrationMethod* for the initialisation and affine registration and *antsRegistration* for the deformable registration. The final result, shown in Figure 3.7, shows a good ventricle alignment without significant skull deformation. It is, however, still possible to notice some diffuse hyper-intense areas adjacent to the inner border of the skull.

3.3.4 CT template integration

3.3.4.1 Registration of native CT scans to MNI space

Figure 3.8 shows three qualitative results from Dataset 1, obtained from the direct affine registration of native CT scans to the CT-MNI template. The examples presented in Figure 3.8 are three of the best
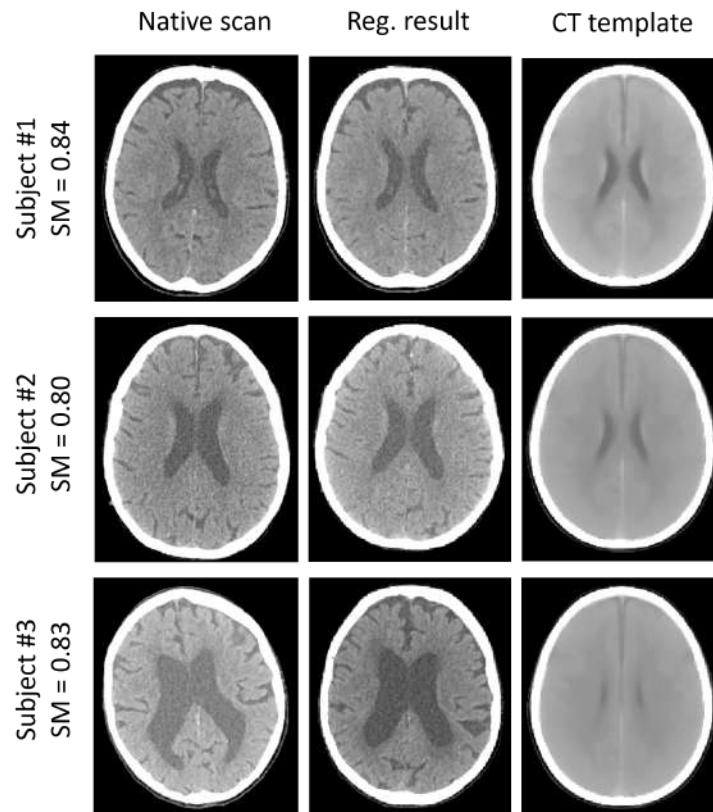
Figure 3.8: Three qualitative registration results from Dataset 1. Direct affine registration of native CT scans to CT-MNI template.

results obtained, according to our visual inspection of images and the registration SM values. The first result (Subj. 1) is moderately acceptable, however we can observe that a successful registration of soft tissue cannot be achieved in scans with larger ventricles using this approach, as it is the case of subjects 2 and 3.

Given the unsatisfactory results of the direct affine registration, a two-step registration approach had to be employed, initialised by a rigid+affine registration of native scans to the study-specific CT template. Figure 3.9 displays the difference between the SM values obtained with a single run of this rigid+affine registration and with the selection of the best value out of four runs, both on every scan of Dataset 2. Repeated registration resulted in an equal or improved performance for all the scans with an initial SM lower than 0.2. From the 839 scans available, 7 were corrupted, which led to the failure of
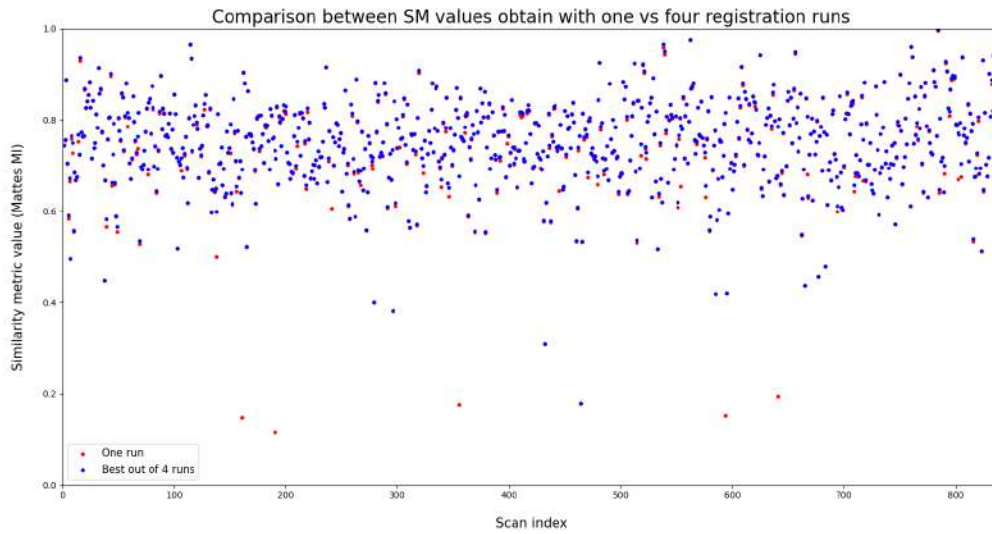
Figure 3.9: Comparison between the SM values obtained with a single run of the rigid+affine registration of native CT scans to the study-specific CT template, and with the selection of the best value out of four runs. Each data point represents one scan, which are identified by their index (x-axis).

their registration. By visually inspecting 160 scans, an SM threshold of 0.65 was empirically defined, below which the result was considered to be sub-optimal. This led to the identification of 96 scans. The correlation between SM values and lesion volume was almost null but significant at a 10% significance level (Pearson $\rho = 0.063$, p=0.07). It was therefore decided not to apply lesion masking to improve registration performance. The plot displayed in Figure 3.10 helps visualise the conclusions stated, i.e., the number of scans identified as sub-optimal and the lack of correlation between lesion volume and SM values, through the distribution of the red data points, which identify the scans with a lesion volume higher than 50mL.

Figure 3.11 presents qualitative results for three scans from Dataset 2, showing three different values of SM, i.e., three different levels of registration success. As this mapping is based on affine registration, it is not possible to achieve great soft-tissue alignment, particularly of the ventricles. Figure 3.12 shows the qualitative registration results of the same three subjects presented in Figure 3.8, this time resulting from the affine registration of native scans to the study-specific CT template. It is possible to observe that, similarly to the results obtained with a direct affine registration of native scans to the CT-MNI template,
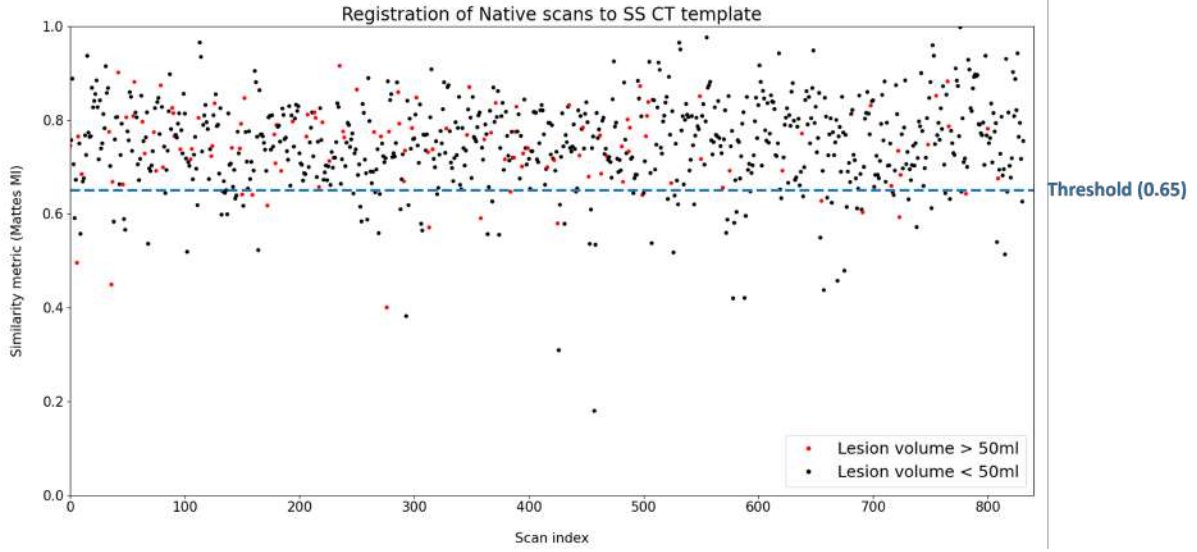
Figure 3.10: Best SM values obtained during the registration of each native scan to the study-specific CT template. Each data point represents one scan, which is identified by its index (x-axis). The result of every scan with an SM value lower than 0.65 (i.e., which point is under the threshold line) is considered to be sub-optimal.

we could only achieve good alignment of the skull border, with a poor soft-tissue alignment (particularly, of the ventricles) on scans with larger ventricles. This was expected from the application of an affine transform. We can nonetheless notice that, due to the initial discrepancy in ventricle size between the study-specific CT template and the CT-MNI template, the registration result presented in Figure 3.12 for each subject is better than the corresponding result in Figure 3.8, which is further confirmed by the increase of the SM value.

3.3.4.2 Lesion localisation

Figure 3.13 shows qualitative results of the atlas mapping for six patients from Dataset 2. Once the back-projected atlas into subject-specific space is obtained, the volumes of lesion per region are calculated through the overlap between the atlas and the lesion segmentation, in this case, the prediction map from BLAST-CT. It is possible to observe that this methodology does not achieve voxel-wise accuracy. Nonetheless, the region atlas was accurately mapped to native space, even in cases where a part of a particular region is heavily occupied by lesions (e.g. patient #2 in Figure 3.13) or collapsed ventricles (e.g. patient #4 in Figure 3.13). Further analysis of both the localisation results and the performance of BLAST-CT are presented in Chapter 4.

3.4 Discussion

3.4.1 CT template construction

The CT template was necessary in order to create an intermediate step between native scans and the MNI parcellated atlas, helping overcome the challenges that multi-modal registration presents, particularly when registering a lesioned subject-specific scan to a standardised image.

The methodology used by Monteiro et al. [7] was initialised by the selection of a random scan as the target, to which all remaining scans were registered and averaged in order to create the first template.
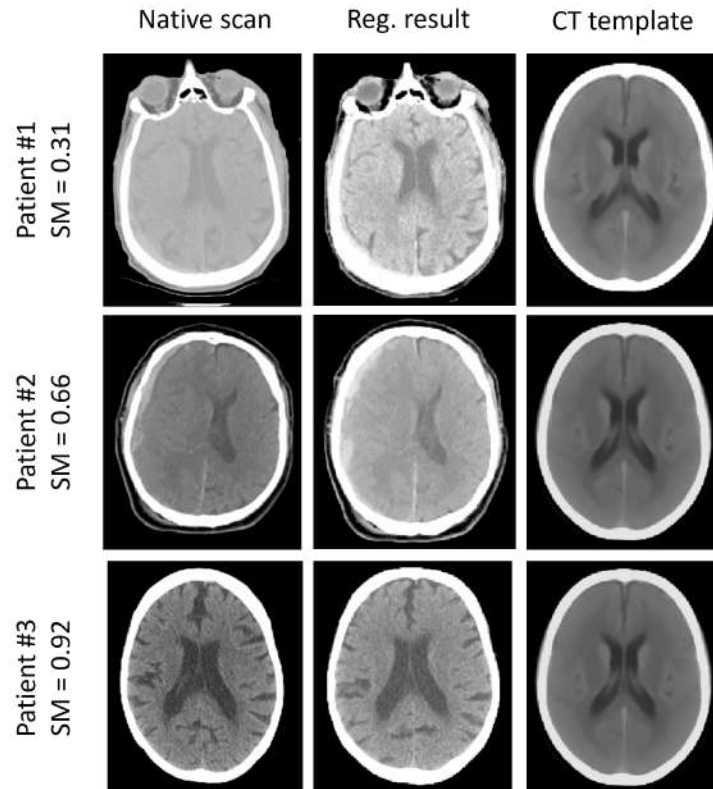
Figure 3.11: Three qualitative registration results from Dataset 2. Affine registration of native scans to the study-specific CT template.

This introduces a bias to the characteristics of that scan, which is removed in our methodology through the use of the standard MRI MNI template as the initial target. Conversely, the use of a higher number of scans would be expected to lead to a blurrier template, due to the consequent increase in averaging. However, the new template is significantly sharper, which further confirms the visual observation that the new methodology employed is more robust and successful. The additional smoothing observed in the previously created template is probably caused by a combination of an insufficient registration performance and the low soft-tissue contrast characteristic of CT. This makes the black structures edges become grey and consequently more diffuse over the registration iterations. An improved registration therefore justifies the increase in the size of structures that can be particularly noticed in the ventricles and cerebellum.

3.4.2 CT template registration to MNI space

Our results show that despite the original high discrepancy in structures sizes (e.g. ventricles), it is possible to accurately register the CT template to the standard MNI MRI template, with good soft-tissue alignment. One registration parameter that was found to have a significant impact on the balance between ventricle and skull deformation was the CC radius used on the deformable registration, where a lower value (4) led to a poor soft-tissue alignment, and a higher value (6) resulted in a good alignment of the ventricles but with high deformation of the skull. 5 was the optimal value found. Despite being able to achieve a good soft-tissue alignment (particularly, of the ventricles) without skull deformation, the mismatch between skull intensities on CT (bright) and MRI (dark), mentioned in Section 2.4, remains. One way to work around this might be to mask the cerebral tissue prior to the registration. This was not however a priority in this dissertation given that, as the atlas mapping to native space is based on affine
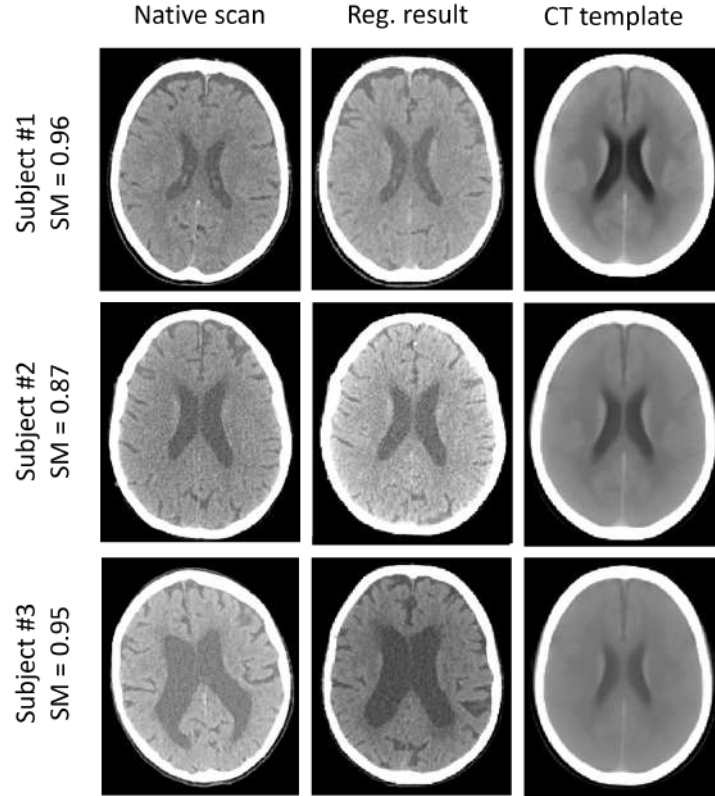
Figure 3.12: Three qualitative registration results from Dataset 1. Affine registration of native scans to the study-specific CT template.

registration, the parcellated regions had to be dilated beyond the skull.

3.4.3 CT template integration

In order to map the parcellated atlas to each native scan, a registration framework between the two spaces had to be defined, which is meant to be applied every time the user wants to get the localisation of a lesion in a scan. Thus, with simplicity and efficiency in mind, we tried to implement an one-step approach, which consisted of an affine registration between each native scan and the CT-MNI template. As would be expected from the application of only an affine transformation, our results show that this approach does not allow for satisfactory alignment of the ventricles of the two images, and therefore a two-step framework is needed. This consisted of an initial rigid + affine registration of native scans to the CT template and an affine + deformable registration of the CT template to MNI space. As shown in Figure 3.12, it is still not possible to achieve great soft tissue alignment of the native scans with the study-specific CT template, as would be expected from the application of an affine transformation. However, the additional application of the deformable transformation yields good results regarding the atlas mapping to native space. Given that only the affine transformation between native space and study-specific space is computed in the localisation algorithm, this approach allows for computational efficiency while still encompassing the benefits of non-linear registration. This further corroborates the need for a study-specific CT template.

Regarding the registration of each native scan and the study-specific CT template, the comparison made between the SM values of one registration run and the best of four registration runs allows us to verify the conjecture that the worst results obtained after one run (SM<0.2) were due to initialisa-
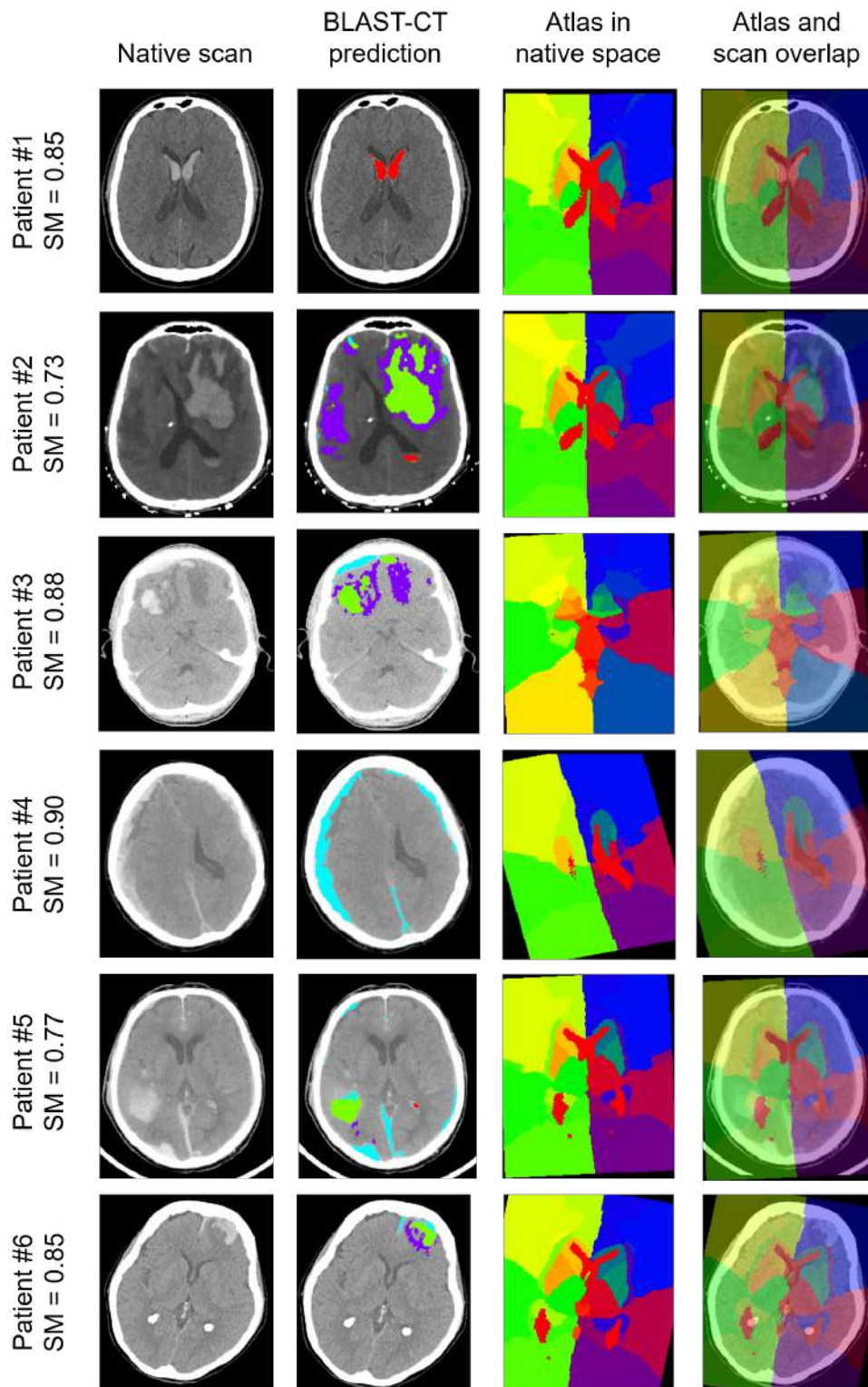
Figure 3.13: Qualitative atlas mapping results from Dataset 2. Images in neurological orientation. Reference segmentation and BLAST-prediction colour legend: Red -IVH; Purple - Oedema; Green - IPH; Light blue - EAH.

3. CT TEMPLATE CONSTRUCTION AND INTEGRATION

tion/sampling errors. Additionally, the lack of correlation between SM values and total lesion volume was not expected, as an increased volume usually leads to a deeper anatomical deformation, which frequently affects registration performance negatively.

Chapter 4

Lesion localisation and BLAST-CT performance analysis

The main lesion localisation output was the volume of each lesion class (IPH, EAH; oedema and IVH) in every atlas region. Additionally, the volume of each atlas region and of the whole brain were also extracted, both in native space. After acquiring these results, it was important to find ways to flag sub-optimal cases, in order to facilitate any future visual inspection or statistical analysis process. An analysis of the lesion class prevalence and of the localisation error per atlas region were also conducted. Finally, the potential bias in the performance of BLAST-CT based on several clinical variables was investigated, contributing to the assessment conducted by Monteiro et al. [7].

4.1 Methods

4.1.1 Lesion localisation error analysis

All analyses described in the present section were completed using the python programming language [116] and the following packages: numpy (numpy.org; last accessed: 31/08/2021), pandas (pandas.pydata.org; last accessed: 31/08/2021), matplotlib (matplotlib.org; last accessed: 31/08/2021) and SimpleITK (simpleitk.org; last accessed: 31/08/2021). FSLeyes (fsl.fmrib.ox.ac.uk; last accessed: 31/08/2021) was used for the visual inspection of images.

4.1.1.1 Sub-optimal results identification

As mentioned in Section 3.3.4.1, after the registration of every native scan of Dataset 2 to the study-specific CT template, an SM (Mattes MI) threshold of 0.65 was empirically defined, below which the result of the registration was considered to be sub-optimal. Two additional flagging methods were conducted after the outcome volumes were calculated:

- Using the volumes of each atlas region, calculated for every patient in native space, 31 box-plots were constructed (one per region) in order to identify outliers. The scans that presented more than five regions [1] with outlier volumes were visually inspected and only the ones showing poor atlas alignment were flagged as sub-optimal.

---

[1] Threshold empirically defined.

4. LESION LOCALISATION AND BLAST-CT PERFORMANCE ANALYSIS

- Any scan with a volume of IVH higher than 1mL localised outside the ventricles was flagged as sub-optimal. This calculation of the excess volume was completed using both the prediction map from BLAST-CT and the reference segmentation maps. The excess volumes found when using the latter were considered to be the baseline errors, being likely due to registration errors and the consequent atlas misalignment.

4.1.1.2 Prevalence brain maps per lesion class

With the goal of understanding which brain regions are more affected by each lesion type, prevalence maps were constructed using the localisation volumes from Dataset 2, excluding the scans previously identified as outliers (Section 4.1.1.1) ($n = n_{localised} - n_{outliers} = 832 - 109 = 723$). These maps are also meant to support the localisation error analysis (Section 4.1.1.3), given that the severity of a specific error rate varies depending on the lesion prevalence in that particular region. Thus, to maximise their accuracy, all prevalence maps were constructed using the volumes per region obtained with the reference segmentation maps.

To create each prevalence map (one per lesion class), a counter is initialised for every atlas region. Going through all the subjects, the counter is incremented by one if the subject has a volume of that lesion class on that region higher than a defined threshold. Two thresholds were applied: 0.1mL and 1mL. The former threshold is set in order to exclude minor misalignments, as the proposed lesion localisation methodology does not allow for voxel-wise accuracy. To facilitate the visualisation of the prevalence values calculated (i.e., the final counter values), these were attributed to each atlas region and visualised as prevalence atlases.

4.1.1.3 False negative and false positive rates maps per lesion class

Two error maps were constructed for each lesion class, with the average values of false negative rate (FNR) and false positive rate (FPR) per atlas region. For each scan, the segmentation map from BLAST-CT was subtracted from the reference segmentation map. The resulting volume corresponds to the false negative map. The inverse operation was then completed to obtain the false positive map. Each true positive volume was in turn computed by subtracting the false positive map to the reference segmentation. Finally, the sum of BLAST-CT segmentation and the false negative map was subtracted from a reference mask in order to obtain the true negative map. In order to normalise this calculation, this reference mask included the full field of view of the parcellated atlas, back-projected into each subject's specific space. All these maps were then overlapped with the parcellated atlas in order to calculate the volume of each variable within each anatomical region and subsequently compute the FNR and FPR values.

These maps were, similarly to the prevalence maps mentioned above, thresholded at 0.1mL, i.e., every scan with a reference volume - or equivalently, a sum of true positive and false negative volumes - lower than 0.1mL was excluded from the average calculation.

Finally, a visual inspection of the scans with the highest total error volume (false negative volume + false positive volume) across all regions was conducted.

4.1.2   Analysis of the BLAST-CT performance bias

All analyses described in this section were completed using the Python programming language [116], with the following packages: pingouin (pingouin-stats.org, last acessed: 15/10/2021), pandas (pandas.pydata.org; last accessed: 31/08/2021) and seaborn (seaborn.pydata.org; last acessed: 15/10/2021).

Performance bias assessment and mitigation in fully automatic DL-based segmentation models are essential steps to ensure an optimal and trustworthy deployment into clinical practice. With the goal of evaluating the effect of several clinical variables in the performance of BLAST-CT, a partial correlation analysis was conducted. The Dice similarity coefficients (DSC) for each lesion class (DSC_{EAH}, DSC_{IPH}, DSC_{Oedema}, DSC_{IVH}), which quantify the agreement between the lesion segmentation by BLAST-CT and the reference, are set as the dependent variables. The clinical variables age, GCS, GOSE, and ISS are set as the independent variables (IVs). The Pearson coefficient was calculated for the tests in which the IV is age, as both this variable and the DSCs are quantitative and continuous. Conversely, for all the remaining tests, the Spearman's correlation coefficient was calculated as GCS, GOSE and ISS can be considered quantitative ordinal variables. For each pairwise coefficient calculation, all the remaining IVs were included as covariates, as well as the time since injury, lesion volume, and the subject's biological sex. Due to the high number of hypothesis tests completed, an adjusted significance level was calculated using the Bonferroni correction.

In order to further investigate the possible sex and age bias effects on the performance of the algorithm, boxplots of the DSC for each lesion class, divided by biological sex and age range, were created. The limits of the three age ranges were defined by the average 33th and 66th percentile between the four lesion classes. Additionally, non-parametric Mann-Whitney U tests were completed to compare the DSC medians between male and female subjects' scans within the same age range, for each lesion class.

4.2   Results

4.2.1   Lesion localisation error analysis

The following sections present the results obtained in the identification of sub-optimal results and in the construction of a prevalence, FNR and FPR map for each lesion class.

4.2.1.1   Sub-optimal results identification

The first method used (i.e. establishing an SM threshold) led to the identification of 96 scans (11.5% of the complete sample) as sub-optimal (Figure 3.10), as the SM resulting from their registration to the study-specific CT template was below the defined threshold (0.65).

Figure 4.1 shows 31 box-plots, each one showcasing the distribution of one atlas region volume. 17 scans had at least 5 regions with outlier volumes, however only 12 scans (1.4% of the complete sample) were identified as outliers after visual inspection. Figure 4.2 shows three of the initially detected outliers. Patients #1 and #3 are two examples that have several regions with volumes over their corresponding box-plots upper fences (upper outlier). Patient #1 shows both a poor mask and atlas alignment, particularly on the frontal and parietal lobes, and it is also an example of a scan excluded by the first criterion, given

4. LESION LOCALISATION AND BLAST-CT PERFORMANCE ANALYSIS
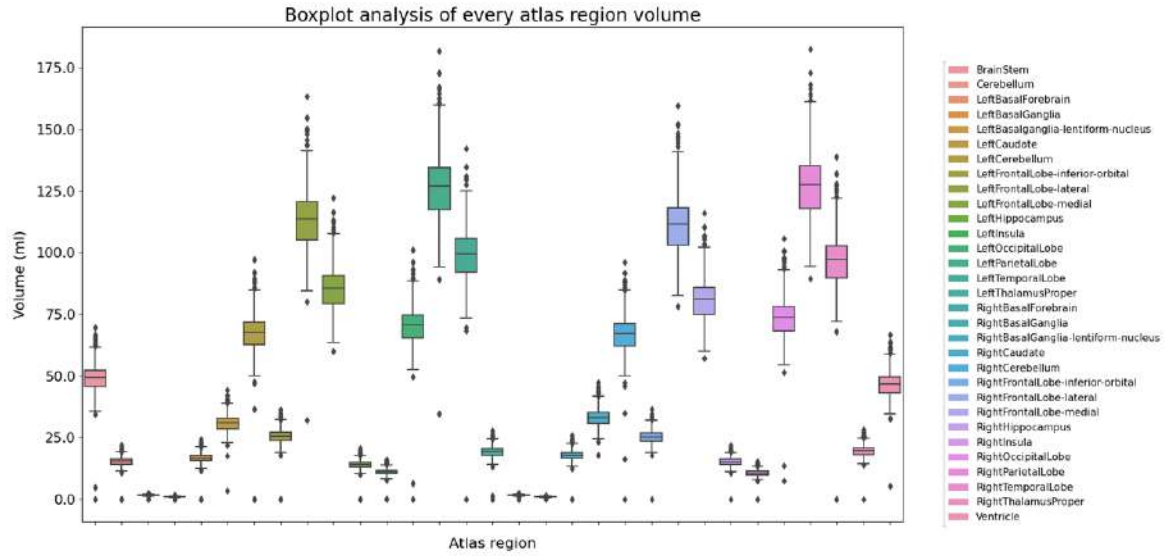
Figure 4.1: Box-plot analysis of the volume of every brain region. The central line of each box-plot represents the median and the box the IQR. The whiskers show the distance of the lowest and highest observed points within the distance of 1.5 times the IQR.

that its registration SM is below 0.65. Conversely, Patient #3 presents an over-segmentation of the brain mask in the occipital lobe and in the cerebellum and brain stem area, and a corresponding poor alignment of the atlas on the same areas, while maintaining a satisfactory alignment of the ventricles and frontal regions. Contrarily, Patient #2 had several regions with volumes under its corresponding box-plot inner fence (inner outlier). Its mask alignment is overall acceptable, with a slight posterior over-segmentation, as well as an under-segmentation of the lower portion of the cerebellum and brainstem. However, the atlas alignment is sufficiently accurate, even in a scan where the ventricles are almost completely suppressed. This was the case for almost every inner outlier, which were not flagged after visual inspection.

Ten scans (1.2% of the complete sample) were flagged with the last method, i.e., had more than 1mL of IVH localised outside of the ventricles, when using the prediction map from BLAST-CT. This inaccurate localisation was consistently caused by a poor alignment of the ventricles in the atlas. After visual inspection of all the identified scans, 3 scans flagged when using the reference segmentation map were observed, which had not been identified when using the prediction map from BLAST-CT, due to under-segmentation of IVH by the algorithm. The Venn diagram in Figure 4.3 shows the overlap between the scans flagged using the three presented approaches.

4.2.1.2 Prevalence brain maps per lesion class

Figures 4.4 and 4.5 show the prevalence maps created with thresholds of 0.1mL and 1mL, respectively. The corresponding tables with the prevalence values per brain region, for each lesion class, are included in Appendix B (Table B.1).

The prevalence of EAH, IPH and oedema lesions is significantly higher in the anterior half of the brain, while IVH lesions are most prevalent in the ventricles. EAH also presents, as expected, higher prevalence in regions contiguous with the cerebral border. We can further notice that only 3 subjects had a volume of IVH higher than 1mL localised in regions outside the ventricles. This indicates that although

36
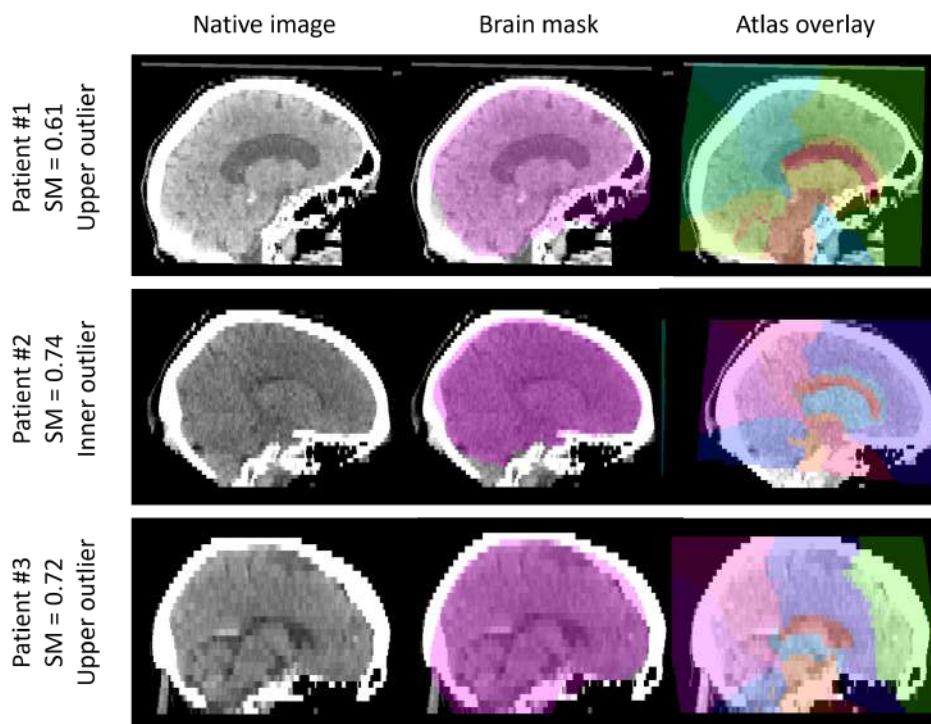
Figure 4.2: Three qualitative examples of scans identified as outliers through the box-plot analysis of brain regions' volumes. Patient #1 shows an overall poor mask and atlas alignment while Patient #3 solely presents mask over-segmentation in the occipital lobe and in the cerebellum and brain stem area. Patient #2 is one of the few examples with several regions with volumes under its corresponding box-plot inner fence (inner outlier). Images in neurological orientation.
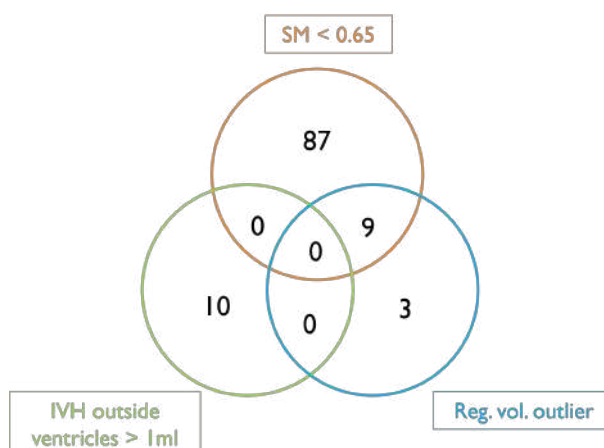
Figure 4.3: Venn diagram of the overlap between the scans that yielded sub-optimal results, identified by three different methods. "SM < 0.65": scans with an SM (Mattes MI) value, resulting from their registration to the study-specific CT template, lower than 0.65; "IVH outside ventricles > 1mL": scans with more than 1mL of IVH localised outside of the ventricles; "Reg. vol. outlier": scans identified as outliers on a box-plot analysis of the volume of every atlas region.

68 patients have a significant volume (>0.1mL) of IVH in the surrounding regions of the ventricles, for 65 patients the incorrectly localised lesion volume ranges between 0.1mL and 1mL.

4.2.1.3 False negative and false positive rates maps per lesion class

Figures 4.6 to 4.9 show the FNR and FPR maps for EAH, IPH, oedema and IVH. Prevalence maps with a threshold of 0.1mL are also included for convenience. The corresponding tables with the FNR and FPR values per brain region, for each lesion class, are included in the Appendix C (Tables C.1, C.2, C.3

4. LESION LOCALISATION AND BLAST-CT PERFORMANCE ANALYSIS
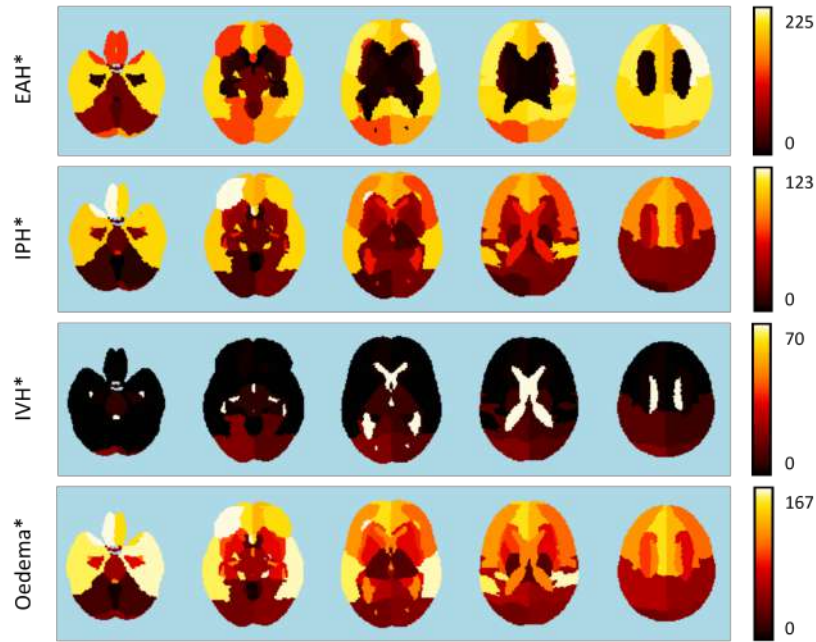
Figure 4.4: Per-class prevalence maps. All maps are displayed in neurological orientation. *Threshold = 0.1 mL.
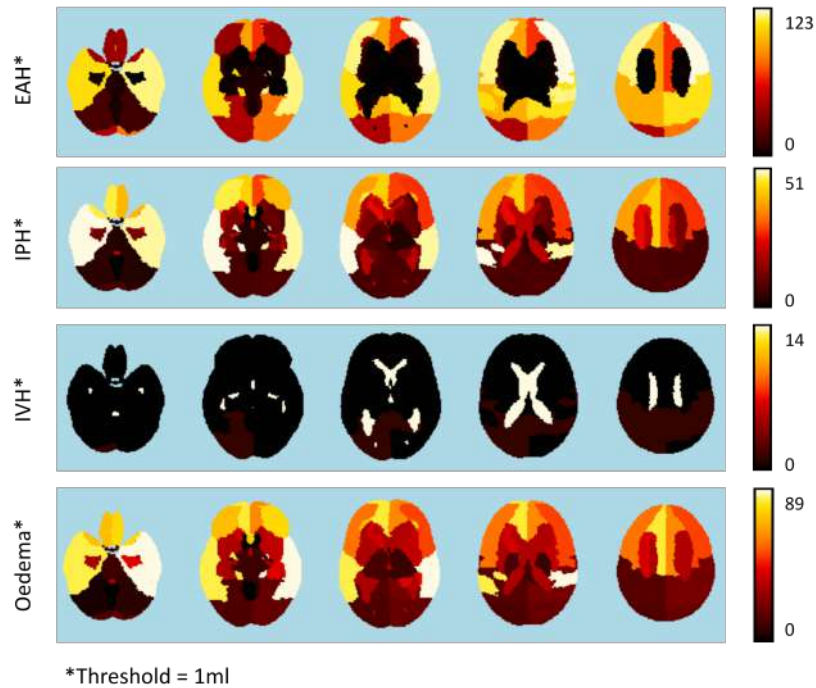
*Threshold = 1ml

Figure 4.5: Per-class prevalence maps. All maps are displayed in neurological orientation. *Threshold = 1 mL.

and C.4). It is possible to notice that the FPR values are consistently lower than the FNR values. This difference is directly correlated with the discrepancy between true lesion volume ("positive" volume, sum of true positive and false negative volume) and true "negative" volume (sum of true negative and false positive volume), denominators of the FNR and FPR formulas, respectively. Therefore, to facilitate the comparison between two rates for the same brain region, the medians of "positive" and "negative" volumes were added to the appended table of each lesion class (Appendix C). Given that the performance of BLAST-CT is positively correlated with lesion volume, the atlas regions were sorted by lesion volume ("positive" volume) in descending order.
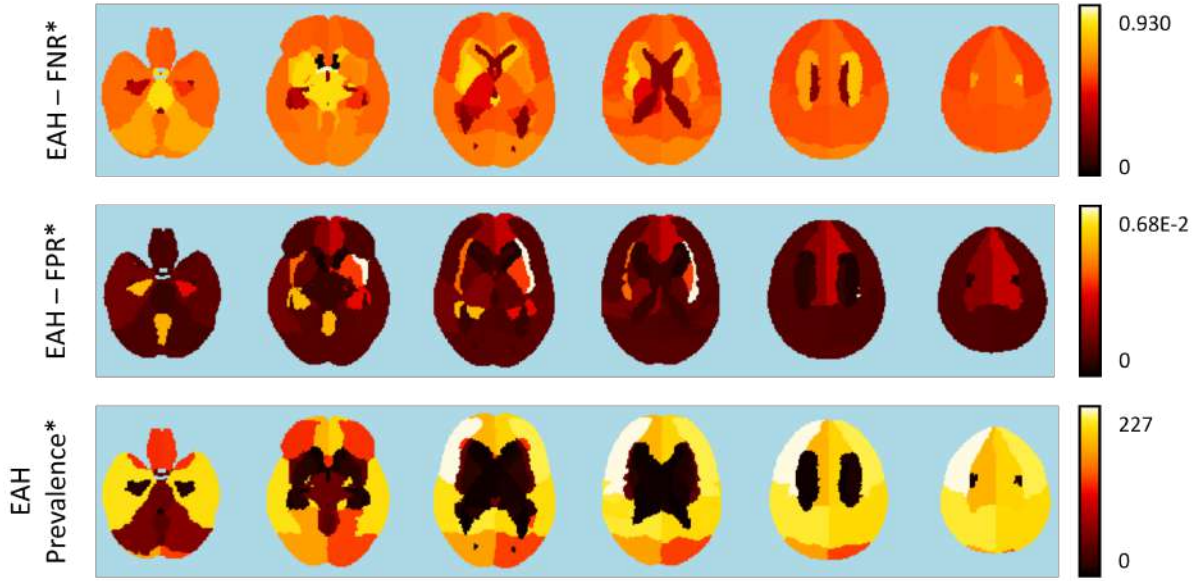
Figure 4.6: False negative rate, false positive rate and prevalence maps for EAH lesions. *Threshold=0.1mL. $n_{Total} = 723$

There is a tendency for higher error rates in regions with lower prevalence and which contain small lesion volumes. The left and right basal ganglia present a FNR higher than 90% for EAH lesions. However, only 6 and 7 scans have an EAH larger than 0.1mL in these regions, respectively, and the average volume within those cases is lower than 0.25mL. Regarding IPH, the left portion of the cerebellum presents the highest FNR, of 76.5%, with a prevalence of 11 scans and a median lesion volume of 0.28mL. Conversely, the highest FPR is in the left basal ganglia, which presents the lowest "negative" volume, of 0.996mL. Among all subjects included in this analysis, no oedema lesion in the cerebellum was correctly segmented and localised, with a FNR of 1. However, only 3 scans have an oedema lesion in this region, with a median volume of 0.22mL. The error analysis of IVH localisation shows a similar situation, with 4 scans presenting a lesion incorrectly localised in the brain stem, with a median volume of 0.24mL, resulting in an average FNR of 97%.

Figure 4.10 shows four qualitative examples of lesion segmentation and atlas mapping, selected based on their high total error volumes from the segmentation of IPH, oedema and EAH lesions. Patient #1 has a total error volume of 33.27mL on the lateral portion of the right frontal lobe, 24.15mL on the right parietal lobe and 19.24mL on the right temporal lobe. In these regions, the FNR/FPR values are 0.69/9.5E-4, 0.97/2.46E-5 and 0.65/1.63E-3, respectively. Regarding Patient #2, the regions with highest total error volume concerning oedema segmentation are the medial portion of the frontal lobe, with 8.70mL and a FNR/FPR pair of 0.67/5.88E-3, and the left temporal lobe, with 10.23mL and a FNR/FPR pair of 0.67/3.57E-4. In regard to IPH lesions, the right caudate presents a total error volume of 5.99mL, with FNR/FPR of 0.54/8.84E-3. The noticeable under-segmentation of oedema in the right temporal lobe in Patient #3 is reflected in a total error volume with 9.17 mL and a corresponding FNR/FPR of 0.96/1.40E-4 in this region. Patient 2, 3 and 4 of Figure 4.10 show a pattern of under-segmented oedema noticed on most of the visually inspected scans.

Additionally, Figure 4.11 shows three qualitative examples with high total error volumes from the segmentation of mainly IVH and IPH lesions. Patient #1 presents an IVH lesion that fully occupies
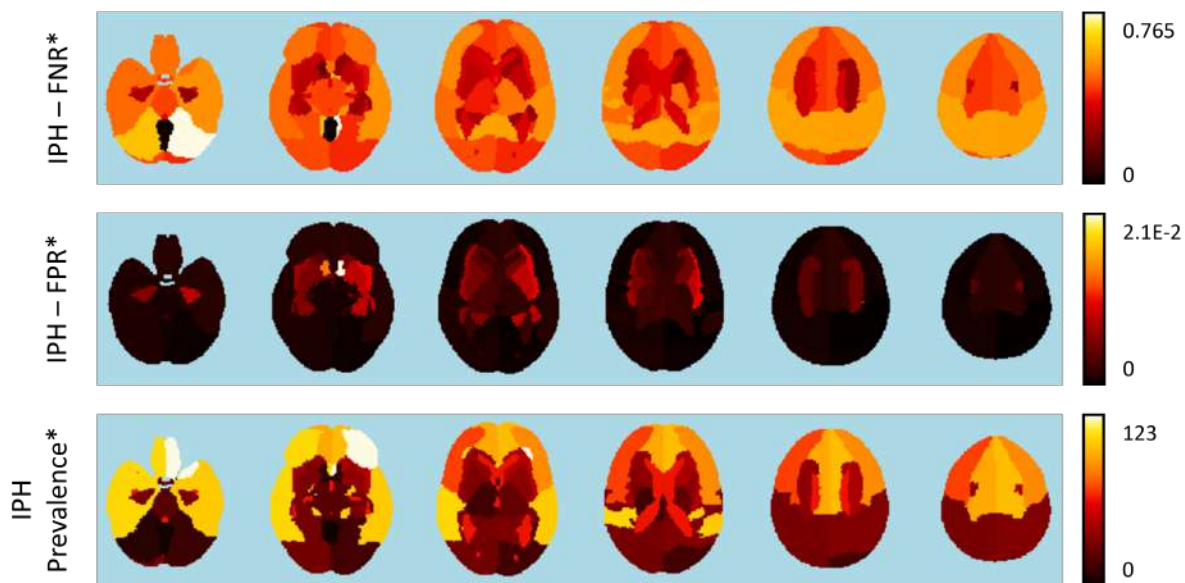
Figure 4.7: False negative rate, false positive rate and prevalence maps for IPH lesions. *Threshold=0.1mL. $n_{Total} = 723$
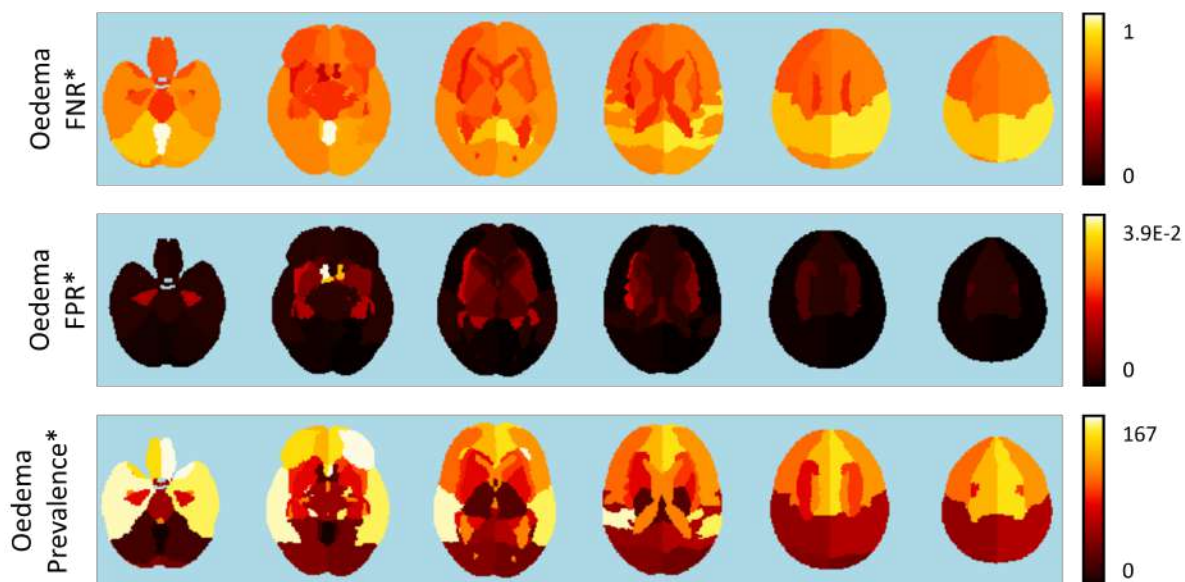
Figure 4.8: False negative rate, false positive rate and prevalence maps for oedema lesions. *Threshold=0.1mL. $n_{Total} = 723$
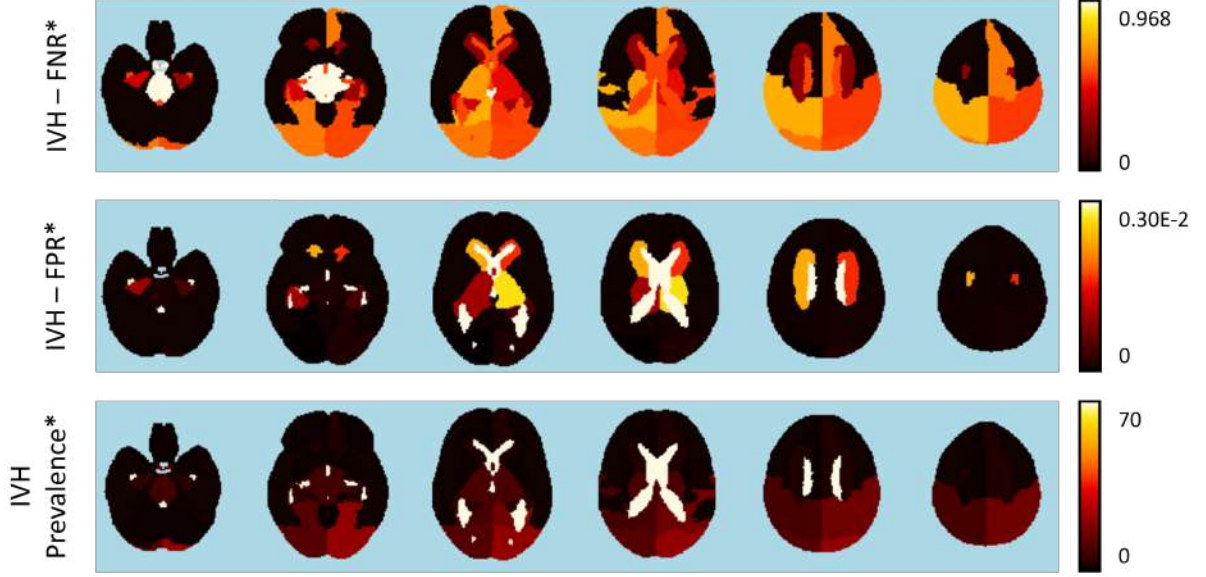
Figure 4.9: False negative rate, false positive rate and prevalence maps for IVH lesions. *Threshold=0.1mL. $n_{Total} = 723$

both ventricles, which was well segmented by BLAST-CT. In this case, the FNR and FPR do not reflect the full extent of the localisation error, as this is solely due to a poor alignment of the ventricles. A considerable under-segmentation of Patient #2's IPH lesion is observable in the posterior portion of the brain, more specifically in the right and central cerebellum and in the right occipital lobe, with a total error volume of 12.65mL, 5.35mL and 4.43mL and a FNR/FPR of 0.55/3.19E-5, 0.95/0.0 and 0.52/2.92E-4, respectively. Regarding their IVH lesion, the highest error volumes are present in the ventricles (3.65mL) and cerebellum (2.73mL). In the latter region, the full predicted lesion is wrongly classified, resulting in a FNR of NaN (no false negative nor true positive volume), and in a high FPR of 0.14. In the former region, the FNR/FPR pair is 0.42/2.31E-2. Finally, Patient #3 presents several extensive IPH and oedema lesions, which severely compromised the structure of the ventricles. Although these lesions seem to be overall well-segmented, the lateral portion of the right frontal lobe presents a slight over-segmentation of oedema, with 9.27mL of total error volume and a FNR/FPR of 0.06/0.012. The IVH lesion within the ventricles was partially classified as IPH, leading to a total error volume of 1.96mL in the ventricles and of 2.56mL in the right parietal lobe, due to a slight misalignment of the atlas. The FNR/FPR pair of these regions is 0.95/4.75E-5 and 0.79/3.39E-5, respectively. Figure 5.1 of the Appendix shows four additional qualitative examples of lesion segmentation and atlas mapping.

4.2.2 Analysis of the BLAST-CT performance bias

Table 4.1 shows the partial correlation coefficients between the segmentation DSCs of each lesion class and several IVs (Age, GCS, GOSE and ISS). Given the initially defined significance level of 0.05 and the total number of statistical tests completed, the Bonferroni adjusted significance level is $\alpha_{new} = \alpha_{initial}/n_{tests} = 0.050/16 = 3.13e^{-3}$. Thus, only the correlation between age and the DSC for EAH lesions ($\rho = 0.193$, p-value < 0.0001) remains statistically significant.

The boxplots in Figure 4.12 show the distribution of the DSC for each lesion class, divided by biological sex and age range. For EAH lesions, the median DSC when including only patients in the age range [6-54] is much lower for female patients (0.13 (Interquartile range (IQR) 0.0-0.45)) than among male patients (0.33 (IQR 0.0-0.58)). Within the age range of [55-69], the median DSC is 0.59 (IQR 0.30-
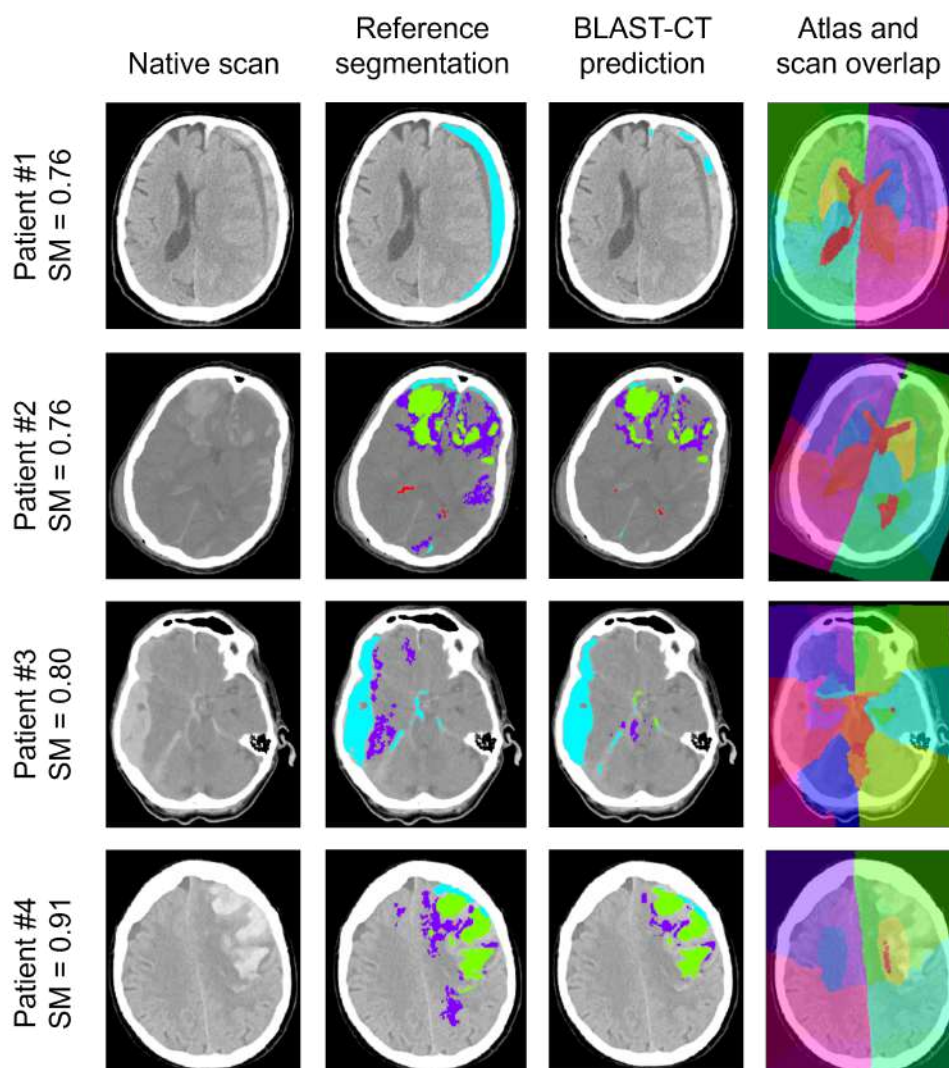
Figure 4.10: Four qualitative segmentation and atlas mapping results. Patient #1: under-segmentation of oedema in left parietal and occipital lobe. Poor mapping of ventricles due to complete collapse of the left ventricle; Patient #2: under-segmentation of oedema, IPH and EAH in the frontal lobe; Patient #3: oedema under-segmentation in the parietal and frontal lobes; Patient #4: under-segmentation of oedema and partial mis-classification of EAH as IPH in the frontal lobe. Reference segmentation and BLAST-prediction colour legend: Red - IVH; Purple - Oedema; Green - IPH; Light blue - EAH. Images in neurological orientation.
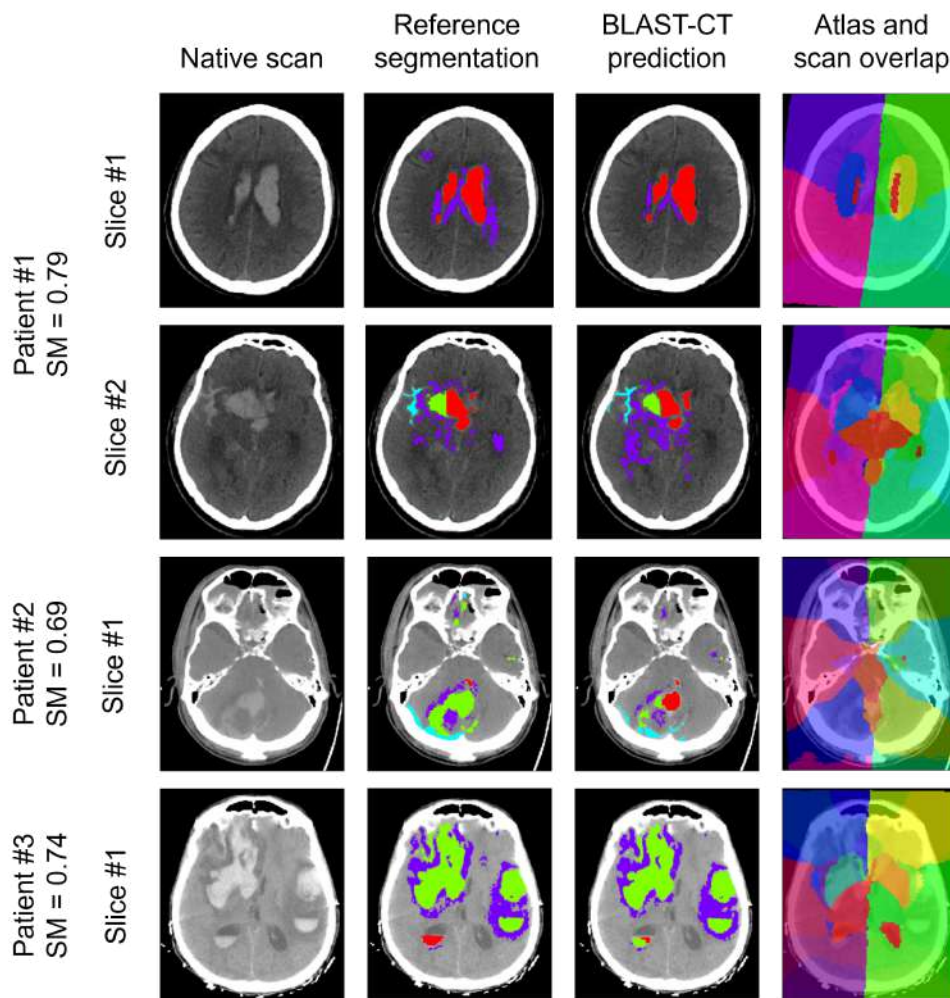
Figure 4.11: Four qualitative segmentation and atlas mapping results. Patient #1: under-segmentation of oedema and IVH localised outside ventricles due to poor mapping of this structure; Patient #2: under-segmentation of IPH (misclassified as IVH), oedema and EAH in the occipital lobe and cerebellum; Patient #3: slight over-segmentation of oedema in the frontal lobe and partial misclassification of an IVH lesion as IPH. Reference segmentation and BLAST-prediction colour legend: Red -IVH; Purple - Oedema; Green - IPH; Light blue - EAH. Images in neurological orientation.

4. LESION LOCALISATION AND BLAST-CT PERFORMANCE ANALYSIS
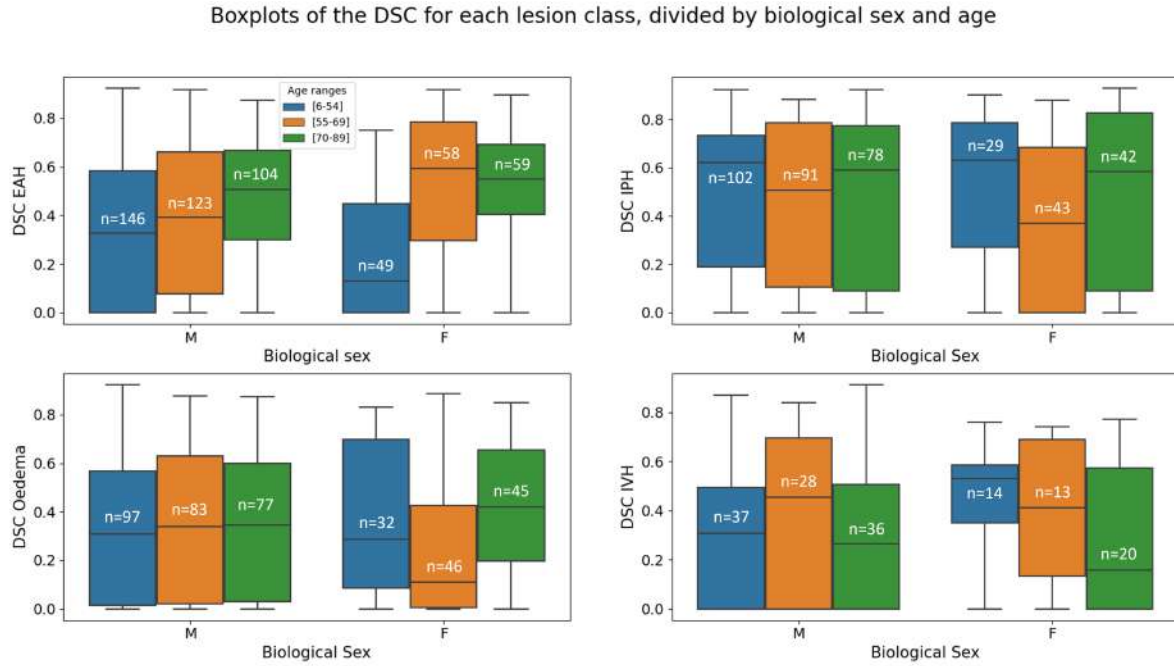
Figure 4.12: Boxplots of DSC per lesion class, biological sex and age, including every scan with either a reference volume higher than 0mL or with a predicted volume higher than 0.1mL. M: Male; F: Female. $n_{Total} = 709$

0.78) and 0.39 (IQR 0.07-0.66) for female and male patients, respectively. The median of male patients increases to 0.51 (IQR 0.30-0.67) within the age range of [70-89]. For female patients of the same age, the median DSC is 0.55 (IQR 0.40-0.69). Regarding IPH lesions, the median DSC for male patients does not present high variance, being 0.62 (IQR 0.19-0.73), 0.50 (IQR 0.10-0.78) and 0.59 (IQR 0.09-0.77) for the age range of [6-54], [55-69] and [70-89], respectively. For female patients, the median DSC of patients within the age ranges of [6-54] (0.63 (IQR 0.27-0.78)) and [70-89] (0.58 (IQR 0.09-0.83)) are similar, decreasing to 0.37 (0.0-0.68) for the age range of [55-69]. Similarly to the previous lesion class, the median DSC for oedema lesions does not vary significantly between age ranges among male patients ([6-54]: 0.31 (IQR 0.01-0.57); [55-69]: 0.34 (IQR 0.02-0.63); [70-89]: 0.35 (IQR 0.03-0.60)). Conversely, for female patients, the low DSC median of 0.11 (IQR 0.005-0.42) for the age range of [55-69] increases to 0.29 (IQR 0.09-0.70) and to 0.42 (IQR 0.20-0.66) for patients within the age ranges of [6-54] and [70-89], respectively. The samples size concerning IVH lesions is considerably lower than in the remaining lesion classes. Within female patients, a steep decrease of the median DSC is observed with increasing age: 0.53 (IQR 0.35-0.58) for [6-54], 0.41 (IQR 0.13-0.69) for [55-69] and 0.16 (IQR 0.0-0.57) for [70-89]. For male patients, patients within the age range of [55-69] present the highest median DSC (0.45 (IQR 0.0-0.69)), followed by the youngest subgroup (0.32 (IQR 0.0-0.49)) and then by the oldest subgroup (0.26 (IQR 0.0-0.50)).

The results of the Mann-Whitney U tests between the DSC of male and female subjects' scans within the same age range, for each lesion class, are presented in Table 4.2. Within the age range of 6 to 54 years old, the difference between the DSC of male and female patients was found to be significant for EAH and IVH lesions, at a significance level of 0.1. For patients in the age range of 55 to 69 years old, the differences between the DSC of male and female patients were found significant for EAH and oedema lesions.

Table 4.1: Partial correlation analysis between segmentation Dice Similarity Coefficient (DSC) (r_s) of each lesion class and several independent variables (IV) (Age, GCS, GOSE and ISS). For each coefficient calculation, all the remaining IV were included as covariates, as well as the time since injury, lesion volume and biological sex. Bold indicates that the correlation is significant at the alpha level corrected by the Bonferroni correction. r_s: Spearman's rank coefficient; ρ: Pearson correlation coefficient. $n_{Total} = 709$. *$p < 0.05$

		EAH	IPH	Oedema	IVH
	n	467	327	330	127
Age	ρ	**0.193**	-0.0839	0.00509	0.0358
	(CI=0.95)	**[0.10, 0.28]**	[-0.19, 0.03]	[-0.10, 0.11]	[-0.15, 0.22]
	p-value	<0.0001	0.1335	0.9275	0.7000
GCS	r_s	0.119*	0.0455	-0.0569	-0.0752
	(CI=0.95)	[0.03, 0.21]	[-0.06, 0.15]	[-0.17, 0.05]	[-0.25, 0.11]
	p-value	0.0107	0.4168	0.3094	0.419
GOSE	r_s	0.0717	0.0902	0.128*	0.0668
	(CI=0.95)	[-0.02, 0.16]	[-0.02, 0.20]	[0.02, 0.23]	[-0.12, 0.24]
	p-value	0.1244	0.1069	0.0215	0.4722
ISS	r_s	-0.00721	0.00263	0.00775	-0.0899
	(CI=0.95)	[-0.10, 0.08]	[-0.11, 0.11]	[-0.1, 0.12]	[-0.27, 0.09]
	p-value	0.8772	0.9625	0.8899	0.3327

Table 4.2: Mann-Whitney U tests between the median DSC male and female subjects' scans within the same age range, for each lesion class. *$p < 0.1$

	EAH	IPH	Oedema	IVH
[6-54]	0.0783*	0.622	0.872	0.0607*
[55-69]	0.0115*	0.146	0.0984*	0.820
[70-89]	0.109	0.831	0.223	0.831

4.3 Discussion

4.3.1 Lesion localisation error analysis

4.3.1.1 Sub-optimal results identification

The definition of an SM threshold for the registration of native scans to the study-specific CT template was found to be the most strict method for identification of sub-optimal results. The low precision of this metric and its lack of physical significance can lead to an erroneous identification of successful results. However, given the high dependence of the mapping results on registration quality, it is nonetheless an efficient way to coarsely flag sub-optimal results for subsequent visual inspection or exclusion prior to statistical analysis.

The visual inspection of every scan with more than five regions with outlier volumes led to the conclusion that these were mainly caused by either a poor atlas alignment, or an over-segmentation of the brain mask (Patients #1 and #3 of Figure 4.2), which leads to an increase in the atlas regions' volumes due to the use of an expanded atlas. Both cases are caused by sub-optimal registration. It is therefore expected that most of the identified scans with this method also present a registration SM lower than 0.65, which was observed (Figure 4.3) and suggests that this approach might be redundant.

Future work might include the definition and implementation of empirical rules regarding the remaining lesion classes. These could encompass the identification of scans with significant volume of EAH not contiguous with the inside limit of the cranial cavity, and within cerebral tissue. However, as the presented localisation method cannot achieve voxel-wise accuracy, considering a discontinuation of one voxel as "not contiguous" might lead to a faulty identification of sub-optimal results. Additionally, given that this class of lesions includes subdural haematomas, extradural haematomas, and traumatic subarachnoid haemorrhages, it is not possible to simply identify scans with a significant volume of EAH outside of a brain mask. IPH lesions would, ideally, also have to be in contact with the skull surface, but in practice this rule is not applicable as patients frequently present IPH lesions with overlying subdural haematomas. A simpler alternative might be to flag the cases with a significant volume of IPH outside the cerebral tissue, i.e., outside a brain mask, or within the ventricles. The same rule can be applied to oedema. One advantage of this approach is the easy availability of the brain mask, which can already be saved to the disk in the localisation code.

4.3.1.2 Prevalence brain maps per lesion class

The prevalence values for each anatomical region and lesion class were computed based on the reference segmentation maps. Thus, all localisation errors causing IVH lesions to be mapped outside of the ventricles were due to misregistration of the atlas. These errors are more prevalent on the posterior part of the brain, namely in the parietal and occipital lobe and in regions contiguous to the ventricles, such as the caudate and thalamus.

As mentioned, the previous empirical definition of a similarity metric threshold required the visual inspection of a considerable amount of scans. This analysis led to the conclusion that ventricles are often over-segmented. Additionally, in some cases, the ventricles are significantly compromised by a lesion, leading to a poor alignment as this displacement cannot be achieved with affine registration, as can be
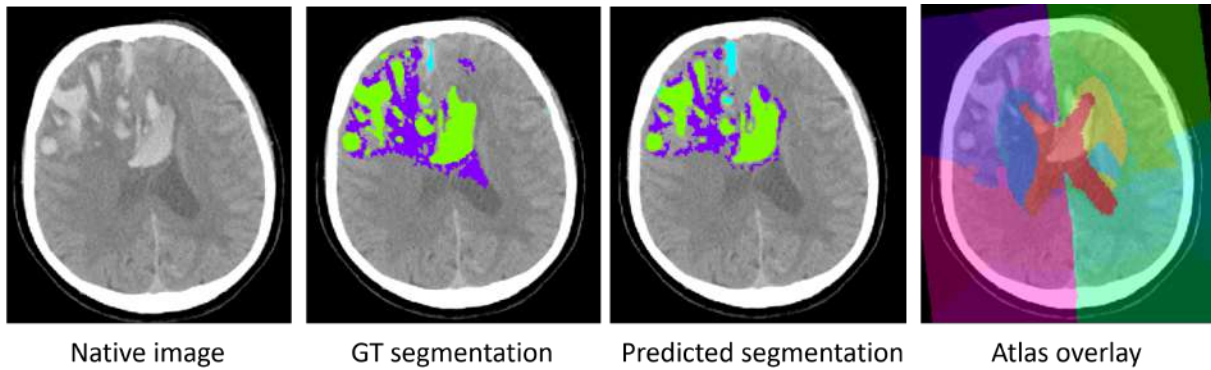
Figure 4.13: Qualitative example of ventricle misregistration. SM = 0.81. IPH: Green; Oedema: Purple; EAH: Blue.

observed in the example displayed in Figure 4.13. These factors present a possible justification for the considerably high number of patients with IPH and oedema lesions localised inside the ventricles.

4.3.1.3   False negative and false positive rates maps per lesion class

The average FNR and FPR was calculated for each anatomical region and lesion class. This analysis presents the limitation of including a high number of variables, given that for each scan that undergoes the lesion localisation process, we extract 124 variables (31 regions * 4 lesion classes). It is therefore challenging to draw any general or summarised conclusions about the error rates. Nonetheless, the presented values are likely to be more valuable when analysed by physicians with deep TBI knowledge, who are able to identify the error in specific regions relevant for the assessment of each lesion class.

As mentioned in Section 3.3.4.1, the quality of the registration between each native scan and the CT template, measured by its SM and on which the atlas mapping accuracy is highly dependent, is not correlated with lesion volume. However, our results indicate that the final localisation error is affected by this variable, as higher FNRs are observed on the regions with lower lesion volumes. This can be due to either a poor segmentation by BLAST-CT, which is known to be influenced by lesion volume, or to the lack of voxel-wise accuracy. With lesions as small as 0.2mL in one anatomical region, a minor registration error may lead to a completely faulty localisation. The high FNR values obtained in regions where the median lesion volume is low (<0.5mL) might indicate that the threshold after which we acknowledge a lesion to be present in a certain region, currently 0.1mL, should be increased.

4.3.2   Analysis of the BLAST-CT performance bias

A partial correlation analysis was conducted between the DSC for each lesion class and several independent variables (age, GCS, GOSE and ISS), also considering the time since injury, lesion volume and biological sex as covariates. Age and biological sex - non-TBI related variables - were included in this assessment as they have been shown to be correlated with TBI outcome, with women generally reporting worse 6-month outcomes, although these differences are dependent on TBI severity and age [117, 118, 119]. Additionally, E Puyol-Antón et al. [120] recently validated the hypothesis that there was a racial bias in DL-based cardiac MR segmentation models as a result of an unbalanced training set. As shown in Table 3.2, the ratio between male and female patients in Dataset 2 - the dataset used to train the CNN of BLAST-CT - is 1.98, which corroborates the need to assess the possible biological sex bias of this algorithm.

4. LESION LOCALISATION AND BLAST-CT PERFORMANCE ANALYSIS

The Bonferroni correction was applied to control for the increase in family-wise error rate across the reported statistical analyses. After this adjustment, only the DSC of EAH lesions and age remain significantly - although weakly - correlated. It is also important to note that the correlation between the DSC of EAH lesions and the GCS, and between the DSC of oedema lesions and the GOSE have been found significant at a 5% significance level. However, given that all the correlation coefficients mentioned are below 0.2, we can conclude that the correlation between each pair of variables, and consequently the performance bias of BLAST-CT, is considerably low.

A higher significance level, of 0.1, was set for the assessment of the differences between the median DSC of male and female subjects within the same age range, as the goal was to further inspect the reason behind the discrepancy in DSC values. Therefore, having a higher probability of committing a type I error is not as concerning. The statistically significant differences in DSC median between the male and female patients within the same age group are likely to be driven by the discrepancies in lesion volume for each one of these sub-groups. Regarding EAH lesions, within the youngest age range ([6-54]), the median lesion volume is 0.89mL (0-27.46)mL among female patients, while it increases to 1.35mL (0-140.2)mL for male patients. For the same lesion class, within the [55-69] age range, the median lesion volume is 4.32mL (0-201.49)mL and 2.28mL (0-179.95)mL for female and male patients, respectively. Additionally, concerning oedema lesions and within the age range of [55-69], male patients present a higher median lesion volume (5.65mL (0-71.54)mL) than female patients (2.96mL (0-75.36)mL). Finally, for IVH lesions of patients within the youngest age range ([6-54]), the median lesion volume is 0.44mL (0-2.79)mL for female patients and 0.175mL (0-12.45)mL for male patients. For every pair-wise comparison mentioned above, an increase in the median lesion volume was accompanied by an increase in the median DSC, which was expected given that the performance of BLAST-CT has been shown to be positively correlated with lesion volumes [7].

Chapter 5

Conclusion and future work

In this dissertation project, a successful registration-based localisation tool for four TBI lesion classes was constructed and tested on a multicentre dataset. A subsequent analysis of lesion prevalence and localisation error rate in each brain region was conducted, along with a bias assessment of the BLAST-CT tool.

The proposed pipeline directly registers a native CT scan to the atlas space, allowing for the back-projection of the parcellated atlas to the subject's specific space. The atlas and each subject's lesion map can then be overlapped and the quantitative volume of lesion within each atlas region calculated. It was shown that, although this pipeline does not allow for voxel-wise precision, it is able to localise lesions accurately even in considerably deformed scans. This tool is currently being deployed[1] and used to process 14,753 scans from the CENTER-TBI study [88]. The output of this process will be fed into relevant clinical research, with the goal of analysing the benefits of lesion localisation quantitative features in combination with full lesion volumes for outcome prediction.

The presented results showed that the localisation error rates calculated for each brain region are overall inversely correlated with lesion volume and prevalence within the dataset. Across all lesion classes, every region with a FNR above 0.75 also presented an average lesion volume below 0.7mL. Furthermore, no considerable bias was found in the performance of BLAST-CT, as all statistically significant correlation coefficients found between DSCs of BLAST-CT and clinical variables (i.e. age, GCS, GOSE and ISS) were below 0.2. The results also suggest that the variation in DSC between male and female patients within a specific age range is not caused by differences in biological sex, but rather by the discrepancy in lesion volume presented by the scans included in each subgroup.

The per-scan output of the localisation tool consists of a large amount of quantitative variables: the total volume of each lesion class (4 variables), the volume of lesion per anatomical region, for each lesion class (31*4=124 variables), and the volume of each anatomical region in native space, as well as of the whole brain (31+1=32 variables). Thus, one important future step to improve the interpretability of these results is to create a patient-specific report that translates the outcomes of our algorithm in a clinically meaningful way. This report should include a quality control section, where the segmentation DSC could be translated into a qualitative metric. To qualitatively evaluate the localisation performance, two metrics could be used: the SM from the registration of the native scan to the study-specific CT tem-

---

[1]The localisation code will be available as part of BLAST-CT, which can be found in https://github.com/biomedia-mira/blast-ct.

5. CONCLUSION AND FUTURE WORK

plate, which reflects how successful the registration was, and the volume of IVH outside the ventricles, since the full volume of intraventricular haemorrhages is expected to be localised inside the ventricles. A combination of these two metrics, possibly translated into a qualitative criterion, could be applied to provide an accessible evaluation of localisation quality. A visual component should also be added, including a colour-coded representation of the lesion map, and of the parcellated atlas alignment in native space. Regarding quantitative information, it would be important to include the total lesion volume per lesion class, the whole brain volume and the volume of each lesion class within the most affected regions, alongside the volume of the region, for reference. Additionally, an "observations" section could be useful to include relevant remarks, such as the volume of IVH outside the ventricles, or if there is an anatomical region that was completely suppressed after the atlas back-projection into native space.

Given the large amount of output features, and the size of the dataset this tool is currently processing, one important part of this project was the identification of scans whose results might have been sub-optimal, as it considerably facilitates the visual inspection process. Three empirical rules were set with this goal in mind: one rule based on the SM from the registration of the native scan to the study-specific CT template, a second rule based on the identification of outliers from box-plots of the volume of each brain region, and a third rule that identified any scans with a significant volume of IVH outside the ventricles. Additional rules should be created concerning EAH, IPH and oedema lesions. A possibly efficient approach could be to identify scans with a considerable volume of IPH and oedema outside a brain mask, or within the ventricles. Conversely, setting a rule for EAH lesions presents several challenges, as this class includes subdural haematomas, extradural haematomas, and traumatic subarachnoid haemorrhages. Thus, it is not possible to simply identify any lesion volume inside the cerebral tissue as incorrect, or scans with lesions not contiguous with the inside limit of the cranial cavity, due to the poor voxel-wise accuracy of the algorithm.

Future work could focus on assessing the segmentation and localisation consistency across serial scans, in order to search and understand any spatial patterns of lesion change. This could be done by selecting a cohort of subjects who have two or more scans, acquired at different time points. Every scan should then be processed by BLAST-CT. Once the volume of lesion per region is calculated, the difference between time points can be assessed.

Bibliography

[1] Michael C Dewan et al. "Estimating the global incidence of traumatic brain injury". In: *Journal of Neurosurgery* 130.4 (2018), pp. 1080–1097.

[2] Dema Najem et al. "Traumatic brain injury: classification, models, and markers". In: *Biochemistry and Cell Biology* 96.4 (2018), pp. 391–406.

[3] Alexandra Brazinova et al. "Epidemiology of traumatic brain injury in Europe: a living systematic review". In: *Journal of Neurotrauma* 38.10 (2021), pp. 1411–1440.

[4] Andrew IR Maas et al. "Traumatic brain injury: integrated approaches to improve prevention, clinical care, and research". In: *The Lancet Neurology* 16.12 (2017), pp. 987–1048.

[5] Stuart Currie et al. "Imaging assessment of traumatic brain injury". In: *Postgraduate Medical Journal* 92.1083 (2016), pp. 41–50.

[6] Peter A Ferrazzano et al. "Use of magnetic resonance imaging in severe pediatric traumatic brain injury: assessment of current practice". In: *Journal of Neurosurgery: Pediatrics* 23.4 (2019), pp. 471–479.

[7] Miguel Monteiro et al. "Multiclass semantic segmentation and quantification of traumatic brain injury lesions on head CT using deep learning: an algorithm development and multicentre validation study". In: *The Lancet Digital Health* 2.6 (2020), pp. 314–322.

[8] Brian H Rowe et al. "Concussion diagnoses among adults presenting to three Canadian emergency departments: missed opportunities". In: *The American Journal of Emergency Medicine* 36.12 (2018), pp. 2144–2151.

[9] Konstantinos Kamnitsas et al. "DeepMedic for brain tumor segmentation". In: *International workshop on Brainlesion: Glioma, multiple sclerosis, stroke and traumatic brain injuries*. Springer. 2016, pp. 138–149.

[10] Marcel Prastawa et al. "A brain tumor segmentation framework based on outlier detection". In: *Medical Image Analysis* 8.3 (2004), pp. 275–283.

[11] Janet M Powell et al. "Accuracy of mild traumatic brain injury diagnosis". In: *Archives of physical medicine and rehabilitation* 89.8 (2008), pp. 1550–1555.

[12] Ilaria Pozzato et al. "Challenges in the acute identification of mild traumatic brain injuries: results from an emergency department surveillance study". In: *BMJ Open* 10.2 (2020).

[13] Wei Li et al. "Traumatic brain injury and age at onset of cognitive impairment in older adults". In: *Journal of Neurology* 263.7 (2016), pp. 1280–1285.

[14] S Fleminger et al. "Head injury as a risk factor for Alzheimer's disease: the evidence 10 years on; a partial replication". In: *Journal of Neurology, Neurosurgery & Psychiatry* 74.7 (2003), pp. 857–862.

BIBLIOGRAPHY

[15] Tanya C Lye and E Arthur Shores. "Traumatic brain injury as a risk factor for Alzheimer's disease: a review". In: *Neuropsychology Review* 10.2 (2000), pp. 115–129.

[16] Corinna Van Den Heuvel, Emma Thornton, and Robert Vink. "Traumatic brain injury and Alzheimer's disease: a review". In: *Progress in Brain Research* 161 (2007), pp. 303–316.

[17] Thamil Mani Sivanandam and Mahendra Kumar Thakur. "Traumatic brain injury: a risk factor for Alzheimer's disease". In: *Neuroscience & Biobehavioral Reviews* 36.5 (2012), pp. 1376–1381.

[18] James F Burke et al. "Traumatic brain injury may be an independent risk factor for stroke". In: *Neurology* 81.1 (2013), pp. 33–39.

[19] Chien-Chang Liao et al. "Stroke risk and outcomes in patients with traumatic brain injury: 2 nationwide studies". In: *Mayo Clinic Proceedings*. Vol. 89. 2. Elsevier. 2014, pp. 163–172.

[20] Stephanie Walsh et al. "A systematic review of the risks factors associated with the onset and natural progression of epilepsy". In: *Neurotoxicology* 61 (2017), pp. 64–77.

[21] Siavash Jafari et al. "Head injury and risk of Parkinson disease: A systematic review and meta-analysis". In: *Movement Disorders* 28.9 (2013), pp. 1222–1229.

[22] Raquel C Gardner et al. "Traumatic brain injury in later life increases risk for P arkinson disease". In: *Annals of Neurology* 77.6 (2015), pp. 987–995.

[23] Paul K Crane et al. "Association of traumatic brain injury with late-life neurodegenerative conditions and neuropathologic findings". In: *JAMA Neurology* 73.9 (2016), pp. 1062–1069.

[24] Pei-Chen Lee et al. "Traumatic brain injury, paraquat exposure, and their relationship to Parkinson disease". In: *Neurology* 79.20 (2012), pp. 2061–2066.

[25] Daniela Impellizzeri et al. "Traumatic brain injury leads to development of Parkinson's disease related pathology in mice". In: *Frontiers in Neuroscience* 10 (2016), p. 458.

[26] Connie Marras et al. "Systematic review of the risk of Parkinson's disease after mild traumatic brain injury: results of the International Collaboration on Mild Traumatic Brain Injury Prognosis". In: *Archives of Physical Medicine and Rehabilitation* 95.3 (2014), S238–S244.

[27] Gentle Sunder Shrestha, Jose I Suarez, and J Claude Hemphill. "Precision medicine in neurocritical care". In: *JAMA Neurology* 75.12 (2018), pp. 1463–1464.

[28] AJ Gardner and R Zafonte. "Neuroepidemiology of traumatic brain injury". In: *Handbook of clinical neurology*. Vol. 138. Elsevier, 2016, pp. 207–223.

[29] Graham Teasdale and Bryan Jennett. "Assessment of coma and impaired consciousness: a practical scale". In: *The Lancet* 304.7872 (1974), pp. 81–84.

[30] Cesar Reis et al. "What's new in traumatic brain injury: Update on tracking, monitoring and treatment". In: *International Journal of Molecular Sciences* 16.6 (2015), pp. 11903–11965.

[31] Graham Teasdale et al. "The Glasgow Coma Scale at 40 years: standing the test of time". In: *The Lancet Neurology* 13.8 (2014), pp. 844–854.

[32] Brandon P Foreman et al. "Usefulness of the abbreviated injury score and the injury severity score in comparison to the Glasgow Coma Scale in predicting outcome after traumatic brain injury". In: *Journal of Trauma and Acute Care Surgery* 62.4 (2007), pp. 946–950.

[33] Bryant Jennett et al. "Disability after severe head injury: observations on the use of the Glasgow Outcome Scale." In: *Journal of Neurology, Neurosurgery & Psychiatry* 44.4 (1981), pp. 285–293.

[34] James M Provenzale. "Imaging of traumatic brain injury: a review of the recent medical literature". In: *American Journal of Roentgenology* 194.1 (2010), pp. 16–19.

[35] Max Wintermark et al. "Imaging evidence and recommendations for traumatic brain injury: conventional neuroimaging techniques". In: *Journal of the American College of Radiology* 12.2 (2015), e1–e14.

[36] David H Livingston et al. "Emergency department discharge of patients with a negative cranial computed tomography scan after minimal head injury". In: *Annals of Surgery* 232.1 (2000), pp. 126–132.

[37] Erskine J Holmes and Rakesh R Misra. *Interpretation of emergency head CT: a practical handbook*. Cambridge University Press, 2017.

[38] Lawrence F Marshall et al. "A new classification of head injury based on computerized tomography". In: *Journal of Neurosurgery* 75.Supplement (1991), S14–S20.

[39] Andrew IR Maas et al. "Prediction of outcome in traumatic brain injury with computed tomographic characteristics: a comparison between the computed tomographic classification and combinations of computed tomographic predictors". In: *Neurosurgery* 57.6 (2005), pp. 1173–1182.

[40] Allen W Brown et al. "Predictive utility of an adapted Marshall head CT classification scheme after traumatic brain injury". In: *Brain Injury* 33.5 (2019), pp. 610–617.

[41] Sunil Munakomi. "A comparative study between Marshall and Rotterdam CT scores in predicting early deaths in patients with traumatic brain injury in a major tertiary care hospital in Nepal". In: *Chinese Journal of Traumatology* 19.1 (2016), pp. 25–27.

[42] Konstantinos Kamnitsas et al. "Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation". In: *Medical Image Analysis* 36 (2017), pp. 61–78.

[43] E Mark Haacke et al. "Common data elements in radiologic imaging of traumatic brain injury". In: *Journal of Magnetic Resonance Imaging* 32.3 (2010), pp. 516–543.

[44] Hamid Reza Talari et al. "The Prognostic Value of Rotterdam Computed Tomography Score in Predicting Early Outcomes Among Children with Traumatic Brain Injury". In: *World Neurosurgery* 125 (2019), e139–e145.

[45] Susan P Baker and Brian O'neill. "The injury severity score: an update". In: *Journal of Trauma and Acute Care Surgery* 16.11 (1976), pp. 882–885.

[46] Benjamin Y Gravesteijn et al. "Toward a New Multi-Dimensional Classification of Traumatic Brain Injury: A Collaborative European NeuroTrauma Effectiveness Research for Traumatic Brain Injury Study". In: *Journal of Neurotrauma* 37.7 (2020), pp. 1002–1010.

[47] Snehashis Roy et al. "Intraparenchymal hemorrhage segmentation from clinical head CT of patients with traumatic brain injury". In: *Medical Imaging 2015: Image Processing*. Vol. 9413. International Society for Optics and Photonics. 2015, pp. 94–130I.

[48] HS Bhadauria, Annapurna Singh, and ML Dewal. "An integrated method for hemorrhage segmentation from brain CT imaging". In: *Computers & Electrical Engineering* 39.5 (2013), pp. 1527–1536.

BIBLIOGRAPHY

[49] W Mimi Diyana W Zaki et al. "Qualitative and quantitative comparisons of haemorrhage intracranial segmentation in CT brain images". In: *TENCON 2011-2011 IEEE Region 10 Conference, Bali, Indonesia*. IEEE. 2011, pp. 369–373.

[50] Haibo Wang, Zhiguo Chen, and Jianzhi Wang. "A segmentation algorithm of intracranial hemorrhage CT image". In: *International Conference on Graphic and Image Processing (ICGIP 2011, Cairo, Egypt)*. Vol. 8285. International Society for Optics and Photonics. 2011, 82856Q.

[51] Bhavna Sharma and K Venugopalan. "Automatic segmentation of brain CT scan image to identify hemorrhages". In: *International Journal of Computer Applications* 40.10 (2012), pp. 0975–8887.

[52] Tao Chan. "Computer aided detection of small acute intracranial hemorrhage on computer tomography of brain". In: *Computerized Medical Imaging and Graphics* 31.4-5 (2007), pp. 285–298.

[53] Bahare Shahangian and Hossein Pourghassem. "Automatic brain hemorrhage segmentation and classification in CT scan images". In: *2013 8th Iranian Conference on Machine Vision and Image Processing (MVIP), Zanjan, Iran*. IEEE. 2013, pp. 467–471.

[54] Esther L Yuh et al. "Computer-aided assessment of head computed tomography (CT) studies in patients with suspected traumatic brain injury". In: *Journal of Neurotrauma* 25.10 (2008), pp. 1163–1172.

[55] Esther L Yuh et al. "Quantitative CT improves outcome prediction in acute traumatic brain injury". In: *Journal of Neurotrauma* 29.5 (2012), pp. 735–746.

[56] Mingjie Sun et al. "Intracranial hemorrhage detection by 3D voxel segmentation on brain CT images". In: *2015 International Conference on Wireless Communications & Signal Processing (WCSP), Nanjing, Jiangsu, China*. IEEE. 2015, pp. 1–5.

[57] Anton Bardera et al. "Semi-automated method for brain hematoma and edema quantification using computed tomography". In: *Computerized Medical Imaging and Graphics* 33.4 (2009), pp. 304–311.

[58] Chun-Chih Liao et al. "Computer-aided diagnosis of intracranial hematoma with brain deformation on computed tomography". In: *Computerized Medical Imaging and Graphics* 34.7 (2010), pp. 563–571.

[59] KN Bhanu Prakash et al. "Segmentation and quantification of intra-ventricular/cerebral hemorrhage in CT scans by modified distance regularized level set evolution technique". In: *International Journal of Computer Assisted Radiology and Surgery* 7.5 (2012), pp. 785–798.

[60] Juha Koikkalainen et al. "Automatic quantification of CT images for traumatic brain injury". In: *2014 IEEE 11th International Symposium on Biomedical Imaging (ISBI), Beijing, China*. IEEE. 2014, pp. 125–128.

[61] Weicheng Kuo et al. "Expert-level detection of acute intracranial hemorrhage on head computed tomography using deep learning". In: *Proceedings of the National Academy of Sciences* 116.45 (2019), pp. 22737–22745.

[62] Geert Litjens et al. "A survey on deep learning in medical image analysis". In: *Medical Image Analysis* 42 (2017), pp. 60–88.

[63] Varun Gulshan et al. "Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs". In: *JAMA* 316.22 (2016), pp. 2402–2410.

[64] Andre Esteva et al. "Dermatologist-level classification of skin cancer with deep neural networks". In: *Nature* 542.7639 (2017), pp. 115–118.

[65] Marios Anthimopoulos et al. "Lung pattern classification for interstitial lung diseases using a deep convolutional neural network". In: *IEEE Transactions on Medical Imaging* 35.5 (2016), pp. 1207–1216.

[66] Jie-Zhi Cheng et al. "Computer-aided diagnosis with deep learning architecture: applications to breast lesions in US images and pulmonary nodules in CT scans". In: *Scientific Reports* 6.1 (2016), pp. 1–13.

[67] Adhish Prasoon et al. "Deep feature learning for knee cartilage segmentation using a triplanar convolutional neural network". In: *International conference on medical image computing and computer-assisted intervention*. Springer. 2013, pp. 246–253.

[68] Mohammad R Arbabshirani et al. "Advanced machine learning in action: identification of intracranial hemorrhage on computed tomography scans of the head with clinical workflow integration". In: *NPJ Digital Medicine* 1.1 (2018), pp. 1–7.

[69] Monika Grewal et al. "Radnet: Radiologist level accuracy using deep learning for hemorrhage detection in ct scans". In: *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), Washington, DC, USA*. IEEE. 2018, pp. 281–284.

[70] Kezia Irene et al. "Segmentation and Approximation of Blood Volume in Intracranial Hemorrhage Patients Based on Computed Tomography Scan Images Using Deep Learning Method". In: *2020 International Workshop on Big Data and Information Security (IWBIS), Depok, Indonesia*. IEEE. 2020, pp. 65–72.

[71] Lu Li et al. "Deep Learning for Hemorrhagic Lesion Detection and Segmentation on Brain CT Images". In: *IEEE Journal of Biomedical and Health Informatics* 25.5 (2020), pp. 1646–1659.

[72] Saurabh Jain et al. "Automatic quantification of computed tomography features in acute traumatic brain injury". In: *Journal of Neurotrauma* 36.11 (2019), pp. 1794–1803.

[73] Sasank Chilamkurthy et al. "Deep learning algorithms for detection of critical findings in head CT scans: a retrospective study". In: *The Lancet* 392.10162 (2018), pp. 2388–2396.

[74] Sasank Chilamkurthy et al. "Development and validation of deep learning algorithms for detection of critical findings in head CT scans". In: *arXiv preprint arXiv:1803.05854* (2018).

[75] Samuel Remedios et al. "Distributed deep learning for robust multi-site segmentation of CT imaging after traumatic brain injury". In: *Medical Imaging 2019: Image Processing*. Vol. 10949. International Society for Optics and Photonics. 2019, 109490A.

[76] Heming Yao et al. "Automated Hematoma Segmentation and Outcome Prediction for Patients with Traumatic Brain Injury". In: *Artificial Intelligence in Medicine* (2020), p. 101910.

[77] Negar Farzaneh et al. "Automated Segmentation and Severity Analysis of Subdural Hematoma for Patients with Traumatic Brain Injuries". In: *Diagnostics* 10.10 (2020), p. 773.

[78] S Lehtonen et al. "Neuropsychological outcome and community re-integration following traumatic brain injury: the impact of frontal and non-frontal lesions". In: *Brain Injury* 19.4 (2005), pp. 239–256.

BIBLIOGRAPHY

[79] Matthew A Warner et al. "Assessing spatial relationships between axonal integrity, regional brain volumes, and neuropsychological outcomes after traumatic axonal injury". In: *Journal of Neurotrauma* 27.12 (2010), pp. 2121–2130.

[80] Ryan M Martin et al. "Traumatic hemorrhagic brain injury: impact of location and resorption on cognitive outcome". In: *Journal of Neurosurgery* 126.3 (2017), pp. 796–804.

[81] Luisa Terroni et al. "Stroke lesion in cortical neural circuits and post-stroke incidence of major depressive episode: a 4-month prospective study". In: *The World Journal of Biological Psychiatry* 12.7 (2011), pp. 539–548.

[82] Cody A Chastain et al. "Predicting outcomes of traumatic brain injury by imaging modality and injury distribution". In: *Journal of Neurotrauma* 26.8 (2009), pp. 1183–1196.

[83] Harri Isokuortti et al. "Characterizing the type and location of intracranial abnormalities in mild traumatic brain injury". In: *Journal of Neurosurgery* 129.6 (2018), pp. 1588–1597.

[84] Marielle Ernst et al. "Association of computed tomography ischemic lesion location with functional outcome in acute large vessel occlusion ischemic stroke". In: *Stroke* 48.9 (2017), pp. 2426–2433.

[85] Melissa Thye and Daniel Mirman. "Relative contributions of lesion location and lesion size to predictions of varied language deficits in post-stroke aphasia". In: *NeuroImage: Clinical* 20 (2018), pp. 1129–1138.

[86] Bram R Diamond et al. "Optimizing the accuracy of cortical volumetric analysis in traumatic brain injury". In: *MethodsX* 7 (2020), p. 100994.

[87] Alina Nickel and Götz Thomalla. "Post-stroke depression: impact of lesion location and methodological limitations—a topical review". In: *Frontiers in Neurology* 8 (2017), p. 498.

[88] Andrew IR Maas et al. "Collaborative European NeuroTrauma effectiveness research in traumatic brain injury (CENTER-TBI) a prospective longitudinal observational study". In: *Neurosurgery* 76.1 (2015), pp. 67–80.

[89] Ivana Despotović, Bart Goossens, and Wilfried Philips. "MRI segmentation of the human brain: challenges, methods, and applications". In: *Computational and Mathematical Methods in Medicine* 2015 (2015).

[90] Francisco PM Oliveira and João RS Tavares. "Medical image registration: a review". In: *Computer Methods in Biomechanics and Biomedical Engineering* 17.2 (2014), pp. 73–93.

[91] Max A Viergever et al. "A survey of medical image registration–under review". In: *Medical Image Analysis* 33 (2016), pp. 140–144.

[92] Fatma El-Zahraa Ahmed El-Gamal, Mohammed Elmogy, and Ahmed Atwan. "Current trends in medical image registration and fusion". In: *Egyptian Informatics Journal* 17.1 (2016), pp. 99–124.

[93] Primoz Markelj et al. "A review of 3D/2D registration methods for image-guided interventions". In: *Medical Image Analysis* 16.3 (2012), pp. 642–661.

[94] Tony CW Mok and Albert Chung. "Fast symmetric diffeomorphic image registration with convolutional neural networks". In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA*. 2020, pp. 4644–4653.

[95] Daniel Rueckert et al. "Nonrigid registration using free-form deformations: application to breast MR images". In: *IEEE Transactions on Medical Imaging* 18.8 (1999), pp. 712–721.

[96] J-P Thirion. "Image matching as a diffusion process: an analogy with Maxwell's demons". In: *Medical Image Analysis* 2.3 (1998), pp. 243–260.

[97] Brian B Avants et al. "Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain". In: *Medical Image Analysis* 12.1 (2008), pp. 26–41.

[98] Arno Klein et al. "Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration". In: *Neuroimage* 46.3 (2009), pp. 786–802.

[99] JB Antoine Maintz and Max A Viergever. "A survey of medical image registration". In: *Medical Image Analysis* 2.1 (1998), pp. 1–36.

[100] William M Wells III et al. "Multi-modal volume registration by maximization of mutual information". In: *Medical Image Analysis* 1.1 (1996), pp. 35–51.

[101] Josien PW Pluim, JB Antoine Maintz, and Max A Viergever. "Mutual-information-based registration of medical images: a survey". In: *IEEE Transactions on Medical Imaging* 22.8 (2003), pp. 986–1004.

[102] Frederik Maes, Dirk Vandermeulen, and Paul Suetens. "Medical image registration using mutual information". In: *Proceedings of the IEEE* 91.10 (2003), pp. 1699–1722.

[103] András P Keszei, Benjamin Berkels, and Thomas M Deserno. "Survey of non-rigid registration tools in medicine". In: *Journal of Digital Imaging* 30.1 (2017), pp. 102–116.

[104] Frederik Maes, Dirk Vandermeulen, and Paul Suetens. "Comparative evaluation of multiresolution optimization strategies for multimodality image registration by maximization of mutual information". In: *Medical Image Analysis* 3.4 (1999), pp. 373–386.

[105] William H Press et al. *Numerical Recipes with Source Code CD-ROM 3rd Edition: The Art of Scientific Computing*. Cambridge University Press, 2007.

[106] Stefan Klein, Marius Staring, and Josien PW Pluim. "Evaluation of optimization methods for nonrigid medical image registration using mutual information and B-splines". In: *IEEE Transactions on Image Processing* 16.12 (2007), pp. 2879–2890.

[107] Guoli Song et al. "A review on medical image registration as an optimization problem". In: *Current Medical Imaging* 13.3 (2017), pp. 274–283.

[108] Brian B Avants et al. "A reproducible evaluation of ANTs similarity metric performance in brain image registration". In: *Neuroimage* 54.3 (2011), pp. 2033–2044.

[109] Michael A Unser, Akram Aldroubi, and Charles R Gerfen. "Multiresolution image registration procedure using spline pyramids". In: *Mathematical Imaging: Wavelet Applications in Signal and Image Processing*. Vol. 2034. International Society for Optics and Photonics. 1993, pp. 160–170.

[110] Erik HW Meijering et al. "Quantitative comparison of sinc-approximating kernels for medical image interpolation". In: *International Conference on Medical Image Computing and Computer-Assisted Intervention, Cambridge, UK* (1999), pp. 210–217.

[111] Jeffrey Tsao. "Interpolation artifacts in multimodality image registration based on maximization of mutual information". In: *IEEE Transactions on Medical Imaging* 22.7 (2003), pp. 854–864.

BIBLIOGRAPHY

[112] ITK. *CenteredTransformInitializerFilter Class Reference*. 2021. URL: https://itk.org/Doxygen/html/classitk_1_1CenteredTransformInitializer.html.

[113] Simple ITK. *Registration Overview*. 2020. URL: https://simpleitk.readthedocs.io/en/master/registrationOverview.html (visited on 09/08/2021).

[114] Jason R Taylor et al. "The Cambridge Centre for Ageing and Neuroscience (Cam-CAN) data repository: Structural and functional MRI, MEG, and cognitive data from a cross-sectional adult lifespan sample". In: *Neuroimage* 144 (2017), pp. 262–269.

[115] Christian Ledig et al. "Robust whole-brain segmentation: application to traumatic brain injury". In: *Medical Image Analysis* 21.1 (2015), pp. 40–58.

[116] Python Software Foundation. *Python*. Version 3.8.3. Feb. 19, 2010. URL: http://www.python.org.

[117] Ana Mikolić et al. "Differences between men and women in treatment and outcome after traumatic brain injury". In: *Journal of Neurotrauma* 38.2 (2021), pp. 235–251.

[118] Ana Mikolic et al. "Explaining outcome differences between men and women following mild traumatic brain injury". In: *Journal of Neurotrauma* ja (2021).

[119] Chantal WPM Hukkelhoven et al. "Patient age and outcome following severe traumatic brain injury: an analysis of 5600 patients". In: *Journal of Neurosurgery* 99.4 (2003), pp. 666–673.

[120] Esther Puyol-Antón et al. "Fairness in cardiac mr image analysis: An investigation of bias due to data imbalance in deep learning based segmentation". In: *International Conference on Medical Image Computing and Computer-Assisted Intervention, Strasbourg, France*. Springer. 2021, pp. 413–423.

Appendices

Appendix A

CT template registration to MNI space

A. CT TEMPLATE REGISTRATION TO MNI SPACE

Table A.1: Fixed parameters of the *antsRegistration* function employed on the non-linear registration of the study-specific CT template to the MRI MNI atlas, during the first phase of iterative parameter search. Three combinations of SMs were tested for the affine and deformable registration: MI or Mattes MI (Mattes) as the SM from both the affine and deformable registration, and MI for affine and Cross Correlation (CC) for deformable registration. SyN: symmetric image normalisation method.

Parameters	Affine registration	Deformable registration (SyN)
Dimensionality	3	3
Interpolation	Welch windowed sinc	Welch windowed sinc
Optimiser	Gradient descent Gradient step = 0.2 —- —-	Gradient descent Gradient step = 0.1 Update field variance in voxel space = 3 Total field variance in voxel space = 0
Histogram matching	No	No
Convergence	Iterations per level: 300 x 200 x 100 Threshold = 10^{-6} Window size = 10	Iterations per level: 400 x 200 x 100 x 50 Threshold = 10^{-6} Window size = 5
Smoothing	4 mm x 2 mm x 1 mm	4 mm x 2 mm x 1 mm x 0 mm
Shrink factors	4 x 2 x 1	6 x 4 x 2 x 1

Table A.2: Parameters of the *antsRegistration* function employed on the non-linear registration of the study-specific CT template to the MRI MNI atlas, after the second phase of iterative parameters optimisation. SyN: symmetric image normalisation method. The definition of a 5-levels pyramidal scheme with an increased number of iterations and heavy initial down-sampling, for deformable registration, should allow for better local alignment while avoiding excessive deformation of the images.

Parameters	Affine registration	Deformable registration (SyN)
Dimensionality	3	3
Similarity metric	Mutual information Metric weight = 1 Nº of bins = 32 Sampling strategy: Regular Sampling percentage = 0.25	Cross correlation Metric weight = 1 Radius = 4 Sampling strategy = Regular Sampling percentage = 0.25
Interpolation	Welch windowed sinc	Welch windowed sinc
Optimiser	Gradient descent Gradient step = 0.2 —- —-	Gradient descent Gradient step = 0.1 Update field variance in voxel space = 3 Total field variance in voxel space = 0
Histogram matching	No	
Convergence	Iterations per level: 300 x 200 x 100 Threshold = 10^{-6} Window size = 10	Iterations per level: 800 x 400 x 200 x 100 x 50 Threshold = 10^{-6} Window size = 5
Smoothing	4 mm x 2 mm x 1 mm	6 mm x 4 mm x 2 mm x 1 mm x 0 mm
Shrink factors	4 x 2 x 1	8 x 6 x 4 x 2 x 1

Appendix B

Prevalence brain maps

B. PREVALENCE BRAIN MAPS

Table B.1: Per-region prevalence values for each lesion type. Prevalence values are obtained by initialising a counter for every atlas region. Going through all the subjects, the counter is increased by 1 if the subject has a volume of that lesion class on that region higher than the defined threshold.

Brain regions	Prevalence (Threshold = 0.1mL)				Prevalence (Threshold = 1mL)			
	EAH	IPH	Oedema	IVH	EAH	IPH	Oedema	IVH
Brain stem	47	20	53	4	4	2	9	0
Cerebellum	36	0	3	0	3	0	0	0
Left basal forebrain	7	6	17	0	0	0	0	0
Left basal ganglia	0	10	29	0	0	0	0	0
Left basal ganglia - lentiform-nucleus	13	36	81	0	1	13	32	0
Left caudate	6	44	93	1	0	25	40	0
Left cerebellum	62	11	19	0	15	3	2	0
Left frontal lobe - inferior-orbital	124	123	167	0	38	48	73	0
Left frontal lobe - lateral	213	86	121	0	119	38	58	0
Left frontal lobe - medial	196	100	151	1	92	44	83	0
Left hippocampus	9	30	74	1	2	11	31	0
Left insula	39	47	99	0	7	18	32	0
Left occipital lobe	128	14	38	18	45	6	12	1
Left parietal lobe	199	32	62	11	96	7	14	1
Left temporal lobe	201	104	159	0	110	51	84	0
Left thalamus proper	5	24	27	6	0	6	11	0
Right basal forebrain	6	4	20	0	0	0	0	0
Right basal ganglia	0	7	25	0	0	0	0	0
Right basal ganglia - lentiform-nucleus	11	31	80	0	1	10	36	0
Right caudate	3	36	79	1	1	12	33	0
Right cerebellum	52	6	18	0	12	2	5	0
Right frontal lobe - inferior-orbital	122	109	149	0	40	41	79	0
Right frontal lobe - lateral	225	70	106	0	123	28	51	0
Right frontal lobe - medial	179	94	133	0	67	29	62	0
Right hippocampus	10	29	80	2	0	12	42	0
Right insula	40	40	88	0	4	7	37	0
Right occipital lobe	168	28	44	11	81	7	17	0
Right parietal lobe	209	30	51	6	112	6	21	1
Right temporal lobe	204	108	165	0	119	50	89	0
Right thalamus proper	3	13	26	6	0	2	8	0
Ventricles	2	63	113	70	0	20	32	14

Appendix C

False negative and false positive rates maps per lesion class

C. FALSE NEGATIVE AND FALSE POSITIVE RATES MAPS PER LESION CLASS

Table C.1: FNR and FPR values per brain region, for EAH lesions. "Positive volume" corresponds to the denominator of the FPR formula, i.e. the sum of the true positive and false negative volumes. "Negative volume" is the denominator of the FNR formula, i.e. sum of the true negative and false positives volumes. Prevalence values (threshold = 0.1mL) included for reference. Rows sorted by "Positive volume", in descending order. An indicative colour scale is used to improve the perception of the lowest and highest FNR and FPR values. Warmer colours are associated with higher FPR and FNR values. Median values are presented for the positive and negative volumes. Average values presented for the FNR and FPR values.

Brain regions	Prevalence	Positive volume (mL)	Negative volume (mL)	FNR	FPR
Left frontal lobe - lateral	213	1.53	611.61	0.52	0.0010
Right temporal lobe	204	1.37	342.77	0.59	0.0015
Right frontal lobe - lateral	225	1.34	591.37	0.53	0.0010
Right parietal lobe	209	1.28	471.51	0.55	0.00095
Left temporal lobe	201	1.22	362.77	0.58	0.0012
Right occipital lobe	168	0.97	327.54	0.61	0.00099
Left parietal lobe	199	0.91	497.96	0.56	0.00086
Left frontal lobe - medial	196	0.89	128.41	0.58	0.0028
Right frontal lobe - medial	179	0.72	127.37	0.56	0.0019
Right caudate	3	0.63	33.04	0.68	0.00035
Left frontal lobe - inferior-orbital	124	0.55	187.32	0.56	0.00090
Left basal ganglia - lentiform-nucleus	13	0.50	16.76	0.63	0.0038
Left occipital lobe	128	0.48	305.24	0.64	0.0011
Right frontal lobe - inferior-orbital	122	0.47	186.46	0.57	0.00088
Left cerebellum	62	0.47	235.62	0.71	0.00074
Ventricles	2	0.38	46.55	0.25	0.00056
Left hippocampus	9	0.37	14.12	0.51	0.0033
Left insula	39	0.33	10.99	0.66	0.0068
Right insula	40	0.32	10.59	0.75	0.0043
Right basal ganglia - lentiform-nucleus	11	0.30	17.87	0.82	0.0012
Right cerebellum	52	0.29	230.87	0.69	0.00089
Brain stem	47	0.26	64.13	0.83	0.00056
Left basal forebrain	7	0.25	1.68	0.90	0.0034
Right hippocampus	10	0.23	15.16	0.36	0.0054
Right thalamus proper	3	0.22	19.58	0.44	0.0017
Left thalamus proper	5	0.21	19.30	0.56	0.0012
Cerebellum	36	0.20	15.54	0.59	0.0053
Right basal forebrain	6	0.15	1.66	0.93	0.00029
Left caudate	6	0.13	30.67	0.77	0.00019
Left basal ganglia	0	0.000	0.99	nan	nan
Right basal ganglia	0	0.000	1.05	nan	nan

Table C.2: FNR and FPR values per brain region, for IPH lesions. "Positive volume" corresponds to the denominator of the FPR formula, i.e. the sum of the true positive and false negative volumes. "Negative volume" is the denominator of the FNR formula, i.e. sum of the true negative and false positives volumes. Prevalence values (threshold = 0.1mL) included for reference. Rows sorted by "Positive volume", in descending order. An indicative colour scale is used to improve the perception of the lowest and highest FNR and FPR values. Warmer colours are associated with higher FPR and FNR values. Median values are presented for the positive and negative volumes. Average values presented for the FNR and FPR values.

Brain regions	Prevalence	Positive volume (mL)	Negative volume (mL)	FNR	FPR
Left caudate	44	1.70	30.49	0.25	0.0046
Left occipital lobe	14	0.92	305.38	0.41	0.00020
Right temporal lobe	108	0.91	343.49	0.49	0.00073
Left temporal lobe	104	0.90	364.15	0.54	0.00091
Left frontal lobe - lateral	86	0.83	613.30	0.50	0.00035
Left hippocampus	30	0.74	14.07	0.24	0.0082
Right frontal lobe - inferior-orbital	109	0.71	186.38	0.49	0.00093
Left frontal lobe - medial	100	0.71	128.58	0.45	0.0012
Left frontal lobe - inferior-orbital	123	0.70	187.17	0.49	0.00093
Right frontal lobe - lateral	70	0.69	591.96	0.53	0.00026
Left basal ganglia - lentiform-nucleus	36	0.65	16.71	0.31	0.0083
Right frontal lobe - medial	94	0.64	127.25	0.42	0.00098
Left insula	47	0.60	10.99	0.38	0.0099
Right basal ganglia - lentiform-nucleus	31	0.56	17.85	0.30	0.0050
Right cerebellum	6	0.54	230.92	0.63	0.00070
Right basal forebrain	4	0.52	1.66	0.25	0.0050
Right caudate	36	0.50	32.99	0.33	0.0035
Left basal ganglia	10	0.50	0.99	0.23	0.021
Right thalamus proper	13	0.48	19.57	0.040	0.0026
Right hippocampus	29	0.46	15.09	0.28	0.0055
Ventricles	63	0.45	46.50	0.35	0.0017
Left thalamus proper	24	0.41	19.28	0.50	0.0014
Right parietal lobe	30	0.41	472.60	0.57	0.00039
Left parietal lobe	32	0.39	498.30	0.57	0.00016
Right occipital lobe	28	0.37	327.80	0.45	0.00057
Right basal ganglia	7	0.37	1.04	0.14	0.014
Right insula	40	0.35	10.58	0.33	0.0069
Brain stem	20	0.34	64.16	0.44	0.00041
Left cerebellum	11	0.28	235.85	0.76	0.00023
Left basal forebrain	6	0.21	1.68	0.65	0.0025
Cerebellum	0	0.000	15.58	nan	nan

C. FALSE NEGATIVE AND FALSE POSITIVE RATES MAPS PER LESION CLASS

Table C.3: FNR and FPR values per brain region, for oedema lesions. "Positive volume" corresponds to the denominator of the FPR formula, i.e. the sum of the true positive and false negative volumes. "Negative volume" is the denominator of the FNR formula, i.e. sum of the true negative and false positives volumes. Prevalence values (threshold = 0.1mL) included for reference. Rows sorted by "Positive volume", in descending order. An indicative colour scale is used to improve the perception of the lowest and highest FNR and FPR values. Warmer colours are associated with higher FPR and FNR values. Median values are presented for the positive and negative volumes. Average values presented for the FNR and FPR values.

Brain regions	Prevalence	Positive volume (mL)	Negative volume (mL)	FNR	FPR
Right temporal lobe	165	1.52	342.77	0.69	0.0016
Left frontal lobe - medial	151	1.19	128.41	0.67	0.0017
Right hippocampus	80	1.16	15.04	0.58	0.014
Left temporal lobe	159	1.12	363.36	0.70	0.0017
Right frontal lobe - inferior-orbital	149	1.08	186.03	0.59	0.0016
Left frontal lobe - lateral	121	0.97	613.12	0.67	0.00029
Right frontal lobe - lateral	106	0.95	591.94	0.59	0.00045
Right frontal lobe - medial	133	0.92	127.06	0.64	0.0020
Left frontal lobe - inferior-orbital	167	0.81	186.95	0.60	0.0012
Left caudate	93	0.78	30.47	0.67	0.0045
Right basal ganglia-lentiform-nucleus	80	0.76	17.75	0.56	0.0099
Right caudate	79	0.76	32.85	0.61	0.0053
Right parietal lobe	51	0.70	472.60	0.81	0.00052
Right occipital lobe	44	0.70	327.80	0.69	0.00054
Right thalamus proper	26	0.70	19.50	0.62	0.0068
Right insula	88	0.68	10.50	0.50	0.017
Left hippocampus	74	0.66	14.01	0.66	0.014
Left basal ganglia - lentiform-nucleus	81	0.64	16.66	0.60	0.0096
Left thalamus proper	27	0.63	19.27	0.69	0.0050
Left insula	99	0.62	10.92	0.56	0.010
Right cerebellum	18	0.56	230.92	0.83	0.00074
Ventricles	113	0.47	46.42	0.54	0.0017
Left occipital lobe	38	0.42	305.38	0.75	0.00030
Brain stem	53	0.38	64.13	0.55	0.0021
Left parietal lobe	62	0.33	498.29	0.92	0.00023
Left cerebellum	19	0.32	235.92	0.78	0.00092
Left basal ganglia	29	0.24	0.99	0.42	0.031
Right basal forebrain	20	0.22	1.65	0.43	0.032
Cerebellum	3	0.22	15.58	1.00	0.00000
Right basal ganglia	25	0.19	1.04	0.43	0.039
Left basal forebrain	17	0.17	1.68	0.67	0.016

Table C.4: FNR and FPR values per brain region, for IVH lesions. "Positive volume" corresponds to the denominator of the FPR formula, i.e. the sum of the true positive and false negative volumes. "Negative volume" is the denominator of the FNR formula, i.e. sum of the true negative and false positives volumes. Prevalence values (threshold = 0.1mL) included for reference. Rows sorted by "Positive volume", in descending order. An indicative colour scale is used to improve the perception of the lowest and highest FNR and FPR values. Warmer colours are associated with higher FPR and FNR values. Median values are presented for the positive and negative volumes. Average values presented for the FNR and FPR values.

| Brain regions	Prevalence	Positive volume (mL)	Negative volume (mL)	FNR	FPR
LeftHippocampus	1	0.56	14.14	0.32	0.00036
Ventricles	70	0.40	46.49	0.56	0.0030
RightCaudate	1	0.37	33.04	0.28	0.0023
LeftThalamusProper	6	0.34	19.30	0.47	0.0027
LeftCaudate	1	0.34	30.67	0.28	0.0016
LeftOccipitalLobe	18	0.28	305.38	0.56	8.50E-05
BrainStem	4	0.24	64.20	0.97	9.15E-05
LeftParietalLobe	11	0.22	498.30	0.55	5.25E-05
RightOccipitalLobe	11	0.21	327.84	0.64	2.27E-05
RightThalamusProper	6	0.16	19.58	0.72	0.0010
RightParietalLobe	6	0.14	472.87	0.75	4.48E-05
RightHippocampus	2	0.12	15.20	0.43	0.00099
LeftFrontalLobe-medial	1	0.12	128.68	0.64	4.34E-05
Cerebellum	0	0.00	15.58	nan	nan
LeftBasalForebrain	0	0.00	1.68	nan	nan
LeftBasalGanglia	0	0.00	1.00	nan	nan
LeftBasalganglia-lentiform-nucleus	0	0.00	16.79	nan	nan
LeftCerebellum	0	0.00	235.94	nan	nan
LeftFrontalLobe-inferior-orbital	0	0.00	187.50	nan	nan
LeftFrontalLobe-lateral	0	0.00	613.43	nan	nan
LeftInsula	0	0.00	11.04	nan	nan
LeftTemporalLobe	0	0.00	365.43	nan	nan
RightBasalForebrain	0	0.00	1.66	nan	nan
RightBasalGanglia	0	0.00	1.05	nan	nan
RightBasalGanglia-lentiform-nucleus	0	0.00	17.89	nan	nan
RightCerebellum	0	0.00	230.92	nan	nan
RightFrontalLobe-inferior-orbital	0	0.00	186.55	nan	nan
RightFrontalLobe-lateral	0	0.00	592.15	nan	nan
RightFrontalLobe-medial	0	0.00	127.40	nan	nan
RightInsula	0	0.00	10.62	nan	nan
RightTemporalLobe	0	0.00	343.60	nan	nan
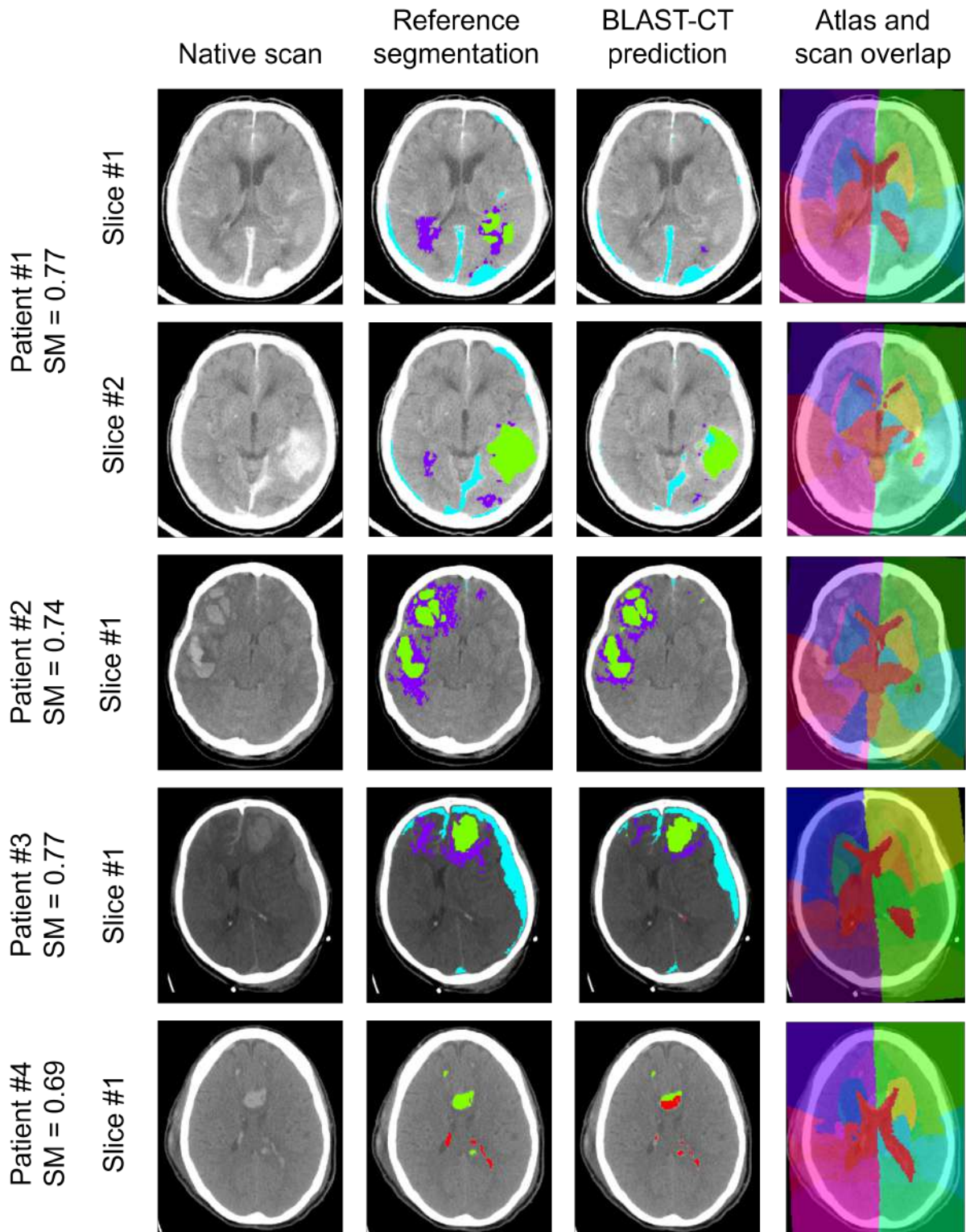
Figure 5.1: Four qualitative examples of lesion segmentation and atlas mapping, selected based on their high total segmentation error volumes. Reference segmentation and BLAST-prediction colour legend: Red -IVH; Purple - Oedema; Green - IPH; Light blue - EAH.