UNIVERSITY OF COPENHAGEN
DEPARTMENT OF MATHEMATICAL SCIENCES

[Master's thesis]

[Name]

[Title of thesis]

[Subtitle]

Date: [June 2021]

Advisor: [Name]

This is a template for the setup itself. The requirements for which information to include varies. You can therefore not simply use the template for your own dissertation. You can get more information about the requirements at your own study office!

Institute:	[Institute for Xxxxx]
Department	[Department for Xxxxx]
Author(s):	[Name]
Title and subtitle:	[Title and subtitle of thesis]
Description:	Lorem ipsum dolor sit amet, consectetur adipiscing elit. Vestibulum porttitor rhoncus dui, vehicula pellentesque odio convallis id. Vestibulum mollis nulla ut sem luctus dictum. Sed scelerisque feugiat arcu, quis faucibus lorem viverra id. Aenean dignissim metus eu ligula sagittis, at commodo magna suscipit.
Advisor:	[Name]
Date:	XX XX XXXX

Table of content

1 Introduction

Introduction

1.1 Background

There are volumetric and microstructural changes in the brain linked to a covid infection, which may result from neuroinflammation and the degeneration of axons [1, 2]. Brian volume changes could also be induced by the lockdown [3].

2 Methods

Methods

2.1 Patient Matching

The following procedure was applied to match the patients:

1. Let us denote the number of 3D T1 MRI-scans (at least 64 slices) for patient k as $n_{\mathrm{S},k}$ and the i'th scan of patient k as $\mathrm{scan}_{k,i}$, $i \in \{1,\, 2,\, ...,\, n_{\mathrm{S},k}\}$. Include patient k for matching if at least two 3D T1 MRI-scans which are sufficiently similar in acquisition parameters, acquired at least 7 days apart, are present. Scans are considered to be sufficiently similar if

$$L_{\mathsf{scan}}(\mathrm{scan}_{k,i},\, \mathrm{scan}_{k,j}) = L_{\mathsf{scan},k,(i,j)} < d_{\mathsf{max}}\,. \tag{1}$$

   $L_{\mathsf{scan},k,(i,j)}$ is defined as the distance between scan i and scan j (of patient k) with $\mathrm{date}(\mathrm{scan}_{k,i}) \neq \mathrm{date}(\mathrm{scan}_{k,j})$. We chose the maximal distance to be $d_{\mathsf{max}} = ...$ The distance $L_{\mathsf{scan},(i,j)}$ is defined as follows:

$$L_{\mathsf{scan},(i,j)} = 3L_{\mathsf{res},(i,j)} + L_{\mathsf{scanner},(i,j)} + L_{\mathsf{B_0},(i,j)}\,. \tag{2}$$

   - The first term measures the difference in resolution:

$$L_{\mathsf{res},(i,j)} = \frac{1}{|R|} \sum_{r \in R} \left(\frac{r_i - r_j}{\mathsf{max}(r_i,\, r_j)} \right)^2, \tag{3}$$

     with $R = \{N_{\mathsf{rows}},\, N_{\mathsf{columns}},\, N_{\mathsf{slices}},\, \Delta d_{\mathsf{rows}},\, \Delta d_{\mathsf{columns}},\, d_{\mathsf{slice}}\}$, see Equation 1 for a detailed description. Since resolution is considered to be crucial to determine the brain volume the term is weighted with a 3.

- The second term $L_{\text{scanner},(i,j)}$ is 0 if the manufacturer model name is the same for both scans and 0 else.
- The third $L_{B_0,(i,j)}$ term contains the difference in the magnetic field strength of the MRI scanner B_0 :

$$L_{B_0,(i,j)} = \left(\frac{B_{0,i} - B_{0,j}}{\max(B_{0,i},\, B_{0,j})} \right)^2 , \tag{4}$$

The standard deviations σ_i , with $i \in r, B_0$ were computed over all the available 3D T1 MRI scans.

Symbol	Unit	Description	dicom tag
N_{rows}	1	Number of Rows	(0028,0010)
N_{columns}	1	Number of Columns	(0028,0011)
N_{slices}	1	Number of Slices	(0054,0081)
Δd_{rows}	mm	Row Spacing	(0028,0030)
$\Delta d_{\text{columns}}$	mm	Column Spacing	(0028,0030)
d_{slice}	mm	Slice Thickness	(0018,0050)

Table 1: This table shows all the resolution parameters. A thorough description of DICOM tags can be found at `https://dicom.innolitics.com/ciods`

2. For convenience we refer from now on to MRI scans that satisfy the criteria specified in 1. simply as scans. We select two groups of patients for our longitudinal study:
   - Patients with at least one scan acquired more than 29 days before the positive COVID diagnosis and at least one scan that was performed after the date of 4 days before the diagnosis. We refer to this group as the positives or positive patients.
   - Potential controls are defined as patients who have no positive COVID test and at least two scans on different days, Or patients with a positive test and at least two scans acquired more than 29 days before the positive COVID diagnosis.
3. Now, we need select a control group from the negatives, which is similar to the positive patients. We match for demographics, comorbidities and minimal difference in time intervals between scans.
   The metric to measure the distance between patients k and l is given by:

$$L_{\text{pat},(k,l)} = 2L_{\text{EHR},(k,l)} + DD_{\text{min},(k,l)} , \tag{5}$$

where the first term includes patient age, sex and comorbidities and is computed as:

$$L_{\text{EHR},(k,l)} = \left(\frac{a_k - a_l}{\sigma_a} \right)^2 \cdot w_a + \left(\frac{b_k - b_l}{\sigma_b} \right)^2 \cdot w_b + \left(\frac{s_k - s_l}{\sigma_s} \right)^2 \cdot w_s + \sum_{c \in C} \left(\frac{c_k - c_l}{\sigma_c} \right)^2 \cdot w_c , \tag{6}$$

with $C = \{\text{Hypertension}, \text{stroke}, ...\}$ c_i is 1 i indicating whether the second term is the minimal difference of time periods between two scans performed on patients k and l. It is computed as follows:

- For a pair of scans (i, j) of patient k that are close enough according to Equation 1, we denote the absolute time interval between scans as: $\Delta t_{(i,j),k} = \text{abs}\Big(\text{date}(\text{scan}_i) - \text{date}(\text{scan}_j)\Big)$.
  Now, to compute the second term in Equation 6 for patients k and l we find two pairs of scans (i, j) and (m, n) of patient k and l respectively that have the minimal difference in time intervals between scans:

$$DD_{\mathsf{min},(k,l)} = \min_{(i,j),(m,n)} \left(\frac{\Delta t_{(i,j),k} - \Delta t_{(m,n),l}}{\sigma_{\Delta t}} \right)^2 \tag{7}$$

To match the patients we select the age on the date of the first scan, if the patient has 2 scans, we select the median age computed over all the scans, excluding the last one, else. We use the following algorithm to perform the matching:

- Given n_{n} negatives and n_{p} positive patients we store pairwise distances defined in Equation 6 in a $n_{\mathsf{n}} \times n_{\text{p}}$ matrix $\mathbf{D}$. The goal now is to find k negatives for every positive patient such that every negative is matched maximally to one positive. This problem can be modeled as as a minimum cost maximum flow problem. To achieve that we create a bipartite graph, see ... . The problem is solved using the simplex algorithm [4, 5] implemented in the python package networkx [6].

  We can look at the pairwise distances as bipartite graph and reducing the minimum weight bipartite matching to a minimum flow max cost problem.

3 [Title 1]

3.1 [Title 2]

References

[1] Yiping Lu, Xuanxuan Li, Daoying Geng, Nan Mei, Pu-Yeh Wu, Chu-Chung Huang, Tianye Jia, Yajing Zhao, Dongdong Wang, Anling Xiao, et al. Cerebral micro-structural changes in covid-19 patients–an mri-based 3-month follow-up study. *EClinicalMedicine*, 25:100484, 2020.

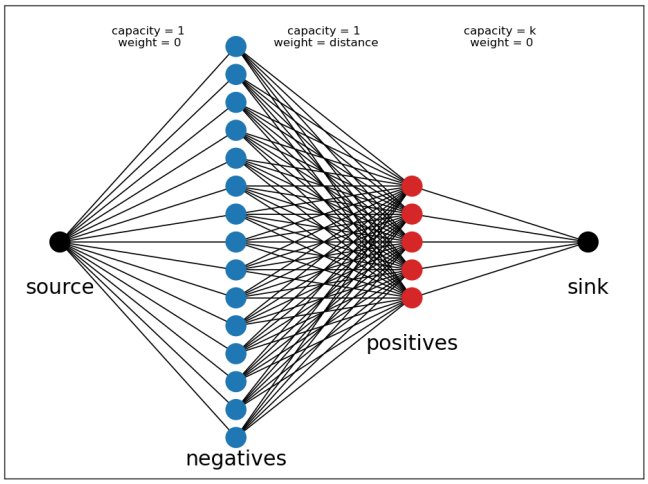

Figure 1: Schematic of the directed graph to model the patient assignment as a max flow min cost problem. The edge capacities from the source to the negative nodes are set to one to prevent double assignment of negatives, weights are set to 0 so we get one unit of flow through each of the edges connecting the source and negatives. To minimize the sum of the distances between the matched patients the edge weight is set to the corresponding distance between the patients. The capacity is set to 1

[2] Elkhonon Goldberg, Kenneth Podell, Daniel K Sodickson, and Els Fieremans. The brain after covid-19: Compensatory neurogenesis or persistent neuroinflammation? *EClinicalMedicine*, 31, 2021.

[3] Tom Salomon, Adi Cohen, Daniel Barazany, Gal Ben-Zvi, Rotem Botvinik-Nezer, Rani Gera, Shiran Oren, Dana Roll, Gal Rozic, Anastasia Saliy, et al. Brain volumetric changes in the general population following the covid-19 outbreak and lockdown. *NeuroImage*, 239:118311, 2021.

[4] Zoltán Király and Péter Kovács. Efficient implementations of minimum-cost flow algorithms. *arXiv preprint arXiv:1207.6381*, 2012.

[5] Richard Barr, Fred Glover, and Darwin Klingman. Enhancements of spanning tree labelling procedures for network optimization. *INFOR: Information Systems and Operational Research*, 17(1):16–34, 1979.

[6] Aric Hagberg, Pieter Swart, and Daniel S Chult. Exploring network structure, dynamics, and function using networkx. Technical report, Los Alamos National Lab.(LANL), Los Alamos, NM (United States), 2008.