A Graph-Theoretic Approach for Analysis of Lesion Changes and Lesions Detection Review in Longitudinal Oncological Imaging

A thesis submitted in fulfillment
of the requirements for
the degree of Master of Science
by

Beniamin Di Veroli

Supervised by
Prof. Leo Joskowicz

The Selim and Rachel Benin
School of Computer Science and Engineering
The Hebrew University of Jerusalem
Jerusalem, Israel

September 2023

Abstract

To assess the status of advanced oncological diseases of a patient, volumetric scans, such as CT or MRI, are acquired periodically. The current clinical protocol provides for lesion detection and measurement in the most recent scan and for the comparison with the measurements of previous scans, i.e., the analysis of lesions changes. This kind of assessment is time-consuming, requires expertise and is subject to observer variability.

This thesis presents a new method and a new workflow for the analysis and review of lesions and volumetric lesion changes in longitudinal scans of a patient. The generic model-based method consists of lesion matching, classification of changes in individual lesions, and detection of patterns of lesions changes. The workflow guides clinicians in the detection of missed and wrongly identified lesions in manual and computed lesion annotations using the analysis of lesion changes. It also serves as heuristic method for the automatic revision of ground truth lesion annotations in longitudinal scans. The tasks are formalized as graph-theoretic problems in which lesions are vertices, edges are lesion pairings computed by overlap-based lesion matching, and changes in individual lesions and of patterns of lesions changes are computed from the properties of the graph and its connected components.

Experimental results on patient studies with ≥ 3 examinations of metastatic lesions in lung (19 patients, 83 CT scans, 1178 lesions), liver (18 patients, 77 CECT scans, 800 lesions) or brain (30 patients, 102 T1W-Gad MRI scans, 317 lesions) patient studies with lesion annotations yielded a high precision 0.92-1.0 and recall 0.91-0.99 for lesion matchings, a high accuracy of 0.87-0.97 for changes in individual lesions and 0.80-0.94 for patterns of lesions changes. The lesion detection review workflow applied to manual and computed lesion annotations yielded 120 and 55 missed lesions and 20 and 164 wrongly identified lesions, for all patient studies, respectively.

Automatic analysis of lesion changes and review of lesion detection in longitudinal studies of oncological patients may improve the accuracy and efficiency of radiological interpretation and disease status evaluation.

Acknowledgements

I thank my supervisor, Prof. Leo Joskowicz, for his academic guidance and support throughout this research, for his patience, and for the opportunity to be part of the CASMIP Lab. I am very grateful for his professional guidance and for the financial support he provided. I feel fortunate to have had such a committed and supportive supervisor.

I also thank Dr. Richard Lederman from the Department of Radiology and Prof. Yigal Shoshan from the Department of Neurosurgery at the Hadassah University Medical Center, my clinical partners for the research, as well as Prof. Jacob Sosna who also supervised this research. I thank Dr. Yossi Srour for facilitating the acquisition and transfer of the data and the manual annotations and for providing feedback throughout the project.

I thank my friends and colleagues at the CASMIP laboratory at the Hebrew University of Jerusalem for helpful discussions, new ideas, and especially for their company and companionship.

Lastly, I would like to thank my dear wife, Dalya Arussy Di Veroli, for the endless support during the past three years. You have a big part of this thesis.

Table of Contents

Chapter 1

Introduction

This thesis presents a novel method for the analysis of lesion changes in time in a series of volumetric medical scans and a workflow to find missed and wrongly identified lesions. The lesion change analysis method is based on lesion matching and handles all kinds of lesion behaviors during the time interval between the acquisitions of the scans, formalized in a graph-theoretic framework. This formalization allows a rigorous matching algorithm evaluation and provides the clinician with a summarized overview of the lesion evolution throughout the patient therapy. The workflow for finding missed and wrongly identified lesions aims to highlight possible errors in lesion detection by focusing clinician attention on the scan regions where unlikely evolution patterns have been identified. Moreover, this workflow can be used as a heuristic automatic tool to correct a series of lesion annotations.

Matching between corresponding lesions and an accurate lesion detection are essential requirement for the automatic generation of a comprehensive report of the most important clinical measurements for evaluating cancer disease progression based on lesion changes analysis. These include both the classification of lesion changes and the quantification of the lesions' changes between the scans. The lesions changes analysis report can be used to quantitatively monitor disease progression, to assess treatment efficacy and to analyze response to therapy of individual lesions.

This chapter presents the clinical motivation for this thesis and an overview of the thesis goals and methodology. Section 1.1 presents the medical background. Section 1.2 explains the importance of radiological lesions follow-up in disease course evaluation. Section 1.3 states the thesis' goals. Section 1.4 presents an overview of the thesis methods. Section 1.5 highlights the novel aspects of the thesis. Section 1.6 describes the thesis organization.

1.1. Background

A lesion is any mass of anomalous tissue in the body. A tumor is any lesion created by abnormal cell growth. Tumors can be either malignant or benign. Malignant tumors are tumors which grow into surrounding areas or spread to additional parts of the body. Cancer is a group of diseases characterized by malignant tumors.

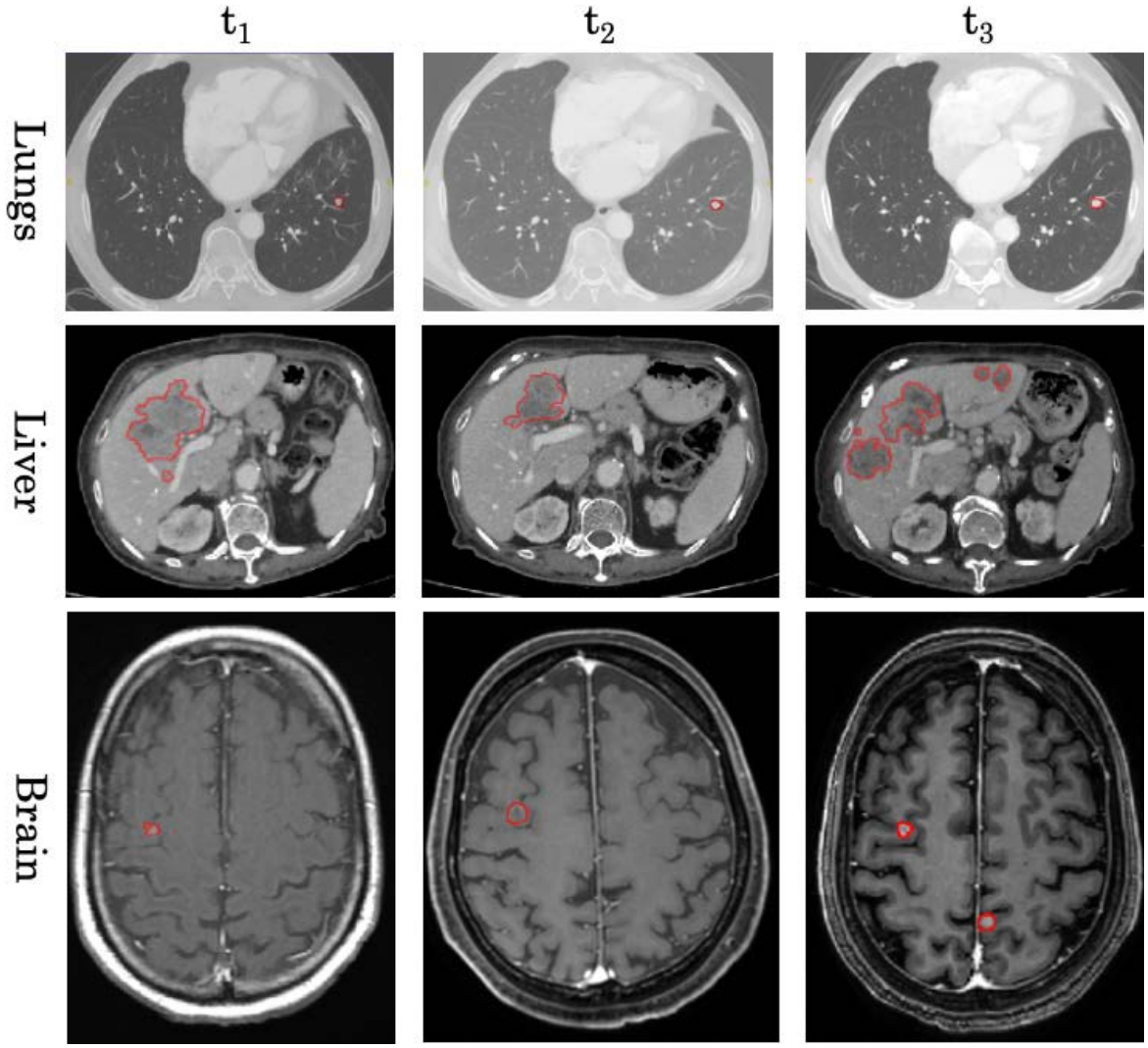


Fig. 1. Example of axial slices of three consecutive scans (taken an t_i) of lungs, liver and brain, containing cancerous lesions. The lungs are scanned with a chest CT, the liver with a contrast-enhanced abdomen CT and the brain with a T1, Gadolinium weighted MRI scan. The lesions in the scans were manually annotated by expert radiologist and their contour is shown in red. Note the different temporal evolution of the lesions in the three examples: for the lungs a finding appear in all the images, consisting in the same lesion. For the liver, the lesion merges and shrinks between t_1 and t_2, and split and expands between t_2 and t_3. In the brain images, the same lesion is persistently seen in all the three scans and a new lesion appears at t_3.

Malignant tumors are classified into two types: primary tumors and secondary tumors, commonly called metastases. The first tumor to form is the primary tumor and is the source of the disease. Cells from the primary tumor may disengage and travel to other parts of the body via the bloodstream, where they may grow into new tumors. The new tumors created by this process are the metastases.

The location and the size of metastases affecting a patient are assessed by acquiring a volumetric image (scan), such as CT or MRI. This thesis deals with metastatic lesions in the lungs, in the liver and in the brain. Fig. 1 shows examples of slices of scans with the annotation of metastatic lesion.

1.2. Radiological follow-up of lesions in longitudinal volumetric scans

Radiological follow-up of oncology patients is an essential component in the evaluation of the disease status, the selection of the treatment, and the response to treatment. CT and MRI scans are routinely acquired every 2-12 months according to the patient's characteristics, disease stage, and treatment regime. As treatments improve, patients live longer, and as the scanners are more available, the number of scans in longitudinal studies increases and their radiological interpretation is more time-consuming and challenging.

Radiological interpretation starts with the detection of cancerous lesions (primary tumors and metastases) in the affected organs in each scan. While most lesions can be detected by examining a single scan, some small and/or subtle or faintly seen lesions require examining the previous and/or following scans to determine if the suspected lesion also appears in them. Fig. 2 illustrates this situation.

Next, analysis of lesion changes is performed. It consists of matching the detected lesions and characterizing their changes over time. Matching two lesions in consecutive scans consists of determining if the lesions appear in the same location in the organ and are thus the same lesion. Changes in individual lesions include changes in the size of persistent lesions, the appearance of new lesions, the disappearance of lesions, and complex lesion changes, e.g., a lesion splitting into several lesions, several lesions merging into one, or the formation of conglomerate lesions. Patterns of lesions changes include lesions that appear in a single scan, lesions that persist in two or more scans, and lesions that split, merge and create complex conglomerates. Identifying the lesions, matching them, and finding the changes in individual lesions and the patterns of lesions changes is challenging because the patient is in a different position each time s/he is scanned,

because the organs undergo deformations and changes in their appearance, and because the scans must be examined simultaneously, side-by-side. Longitudinal reading differs from diagnostic reading since the goal is to find and quantify the differences between the scans, rather than to find abnormalities in a single scan.

In current practice, lesion detection and analysis of lesion changes are partial, approximate and error-prone. The RECIST 1.1 guidelines call for finding new lesions (if any), identifying up to the five largest lesions in each scan in the slice where they appear largest, manually measuring their diameters, and comparing their difference (Eisenhauer et al., 2009). While volumetric measures of individual lesions and of all lesions (tumor burden) have long been established as more accurate and reliable than linear measurements of representative lesions, they are not used clinically because they require manual lesion delineation and lesion matching in unregistered scans, which is time-consuming and subject to variability (Joskowicz et al., 2019).

Corrections made to lesion annotations may alter the disease status as defined by RECIST 1.1 (Joskowicz et al., 2023). Thus, if a lesion was detected in the current scan but was missed in the previous scans, the current scan lesion is identified as new and thus indicates progressive disease. However, if upon review, a missed matching lesion is found in the previous scan, the lesion is persistent and may indicate stable disease.

We hypothesized that the identification of some missed (false negatives) and wrongly identified lesions (false positives) requires the examination of two or more scans and that it is related to suspicious patterns of lesions changes. In previous papers, we showed that a network trained on pairs of CT scans of a patient increases the precision and recall of the detection of liver metastases (Szeskin et al, 2023) and that analysis of lesion changes can be performed by matching pairs of consecutive scans (Rochman et al. 2023).

In this thesis, we extend this approach to longitudinal studies with three or more scans. We define changes in individual lesions and patterns of lesions changes and present a workflow and an automatic method for the review of lesion annotations in longitudinal scans. We show that the detection of some missed and wrongly identified lesions that are hard to find because they require the examination of three or more scans, is related to patterns of lesions changes. Fig. 2 illustrates one such pattern.

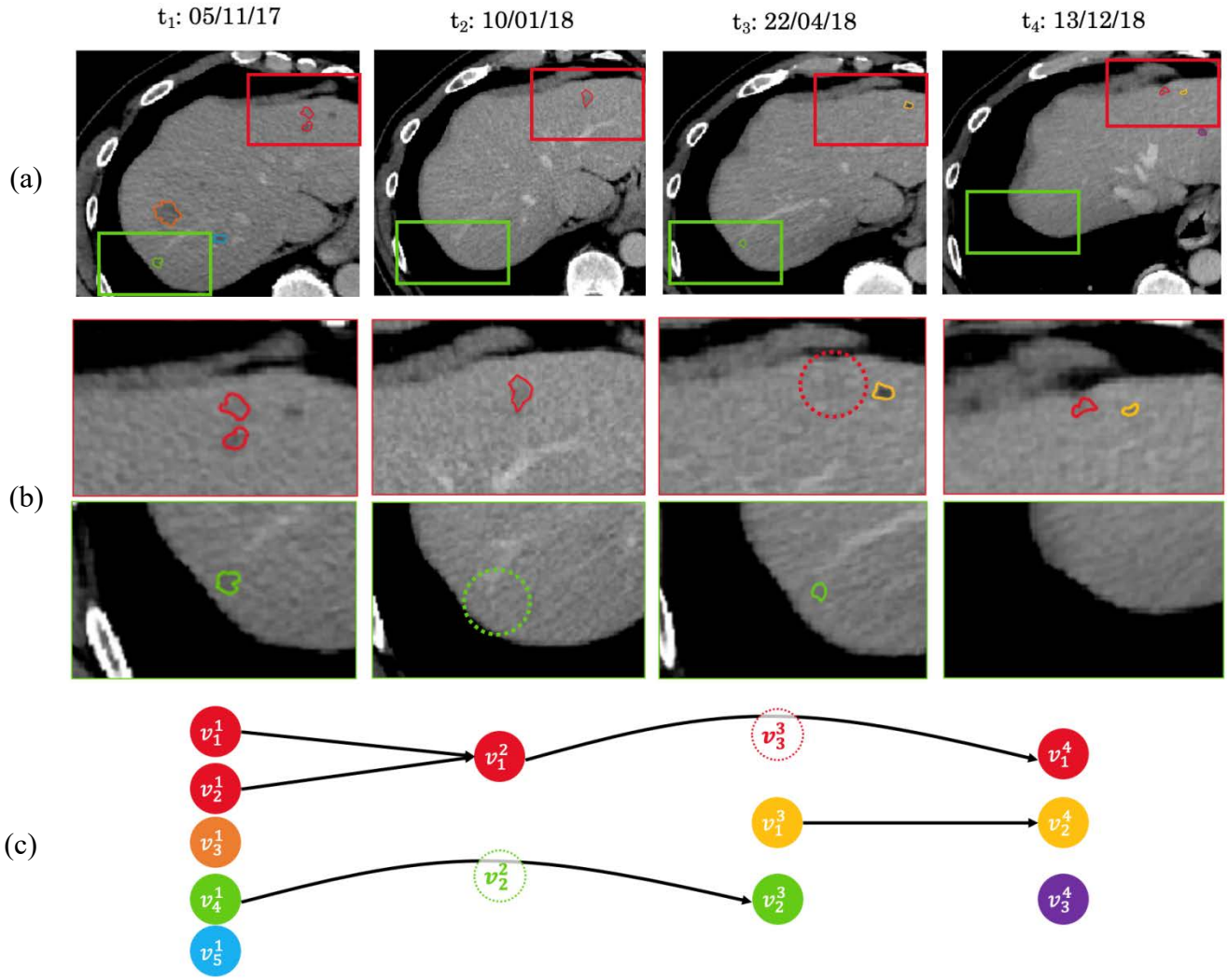


Fig. 2. Example of a longitudinal study of a patient with liver metastases: (a) axial slices from four consecutive unregistered CECT liver scans S_i, acquired at times $t_i, 1 \leq i \leq 4$ with their lesion annotations l_j^i (colored contours) performed by an expert radiologist. Rectangles show two suspected missing lesion regions; (b) details of the two suspected missing lesions regions; dotted circles indicate the suspected missed lesion area; (c) lesion changes graph: vertex v_j^i (solid circle) is associated with lesion l_j^i in scan S_i (same color); edges (v_j^i, v_k^l) (arrows) correspond to matches between lesions in scans S_j, S_k (straight arrows for consecutive edges, arched for non-consecutive edges). Non-consecutive edges indicate possibly missed lesions. Upon review of the non-consecutive edges, two missed lesions are identified in the scans, l_2^2 in S_2 and l_3^3 in S_3, in the area inside the green and dotted circles, respectively, and their corresponding vertices v_2^2 and v_3^3 are added to the lesion changes graph (dotted circles).

1.3. Thesis goals

The main goals of this thesis are to develop a generic method for the accurate and reliable analysis of lesion evolutionary changes in longitudinal series of medical scans and a generic workflow for lesion detection review based on lesion evolutionary patterns. The specific goals are:

1. **Lesion change graph formalization:** describe the lesion changes analysis as a graph-theoretic problem.
2. **Lesion change analysis method**: develop a generic, model-based method for matching corresponding lesions in consecutive and non-consecutive scans and classify individual lesions changes and patterns of lesion changes.
3. **Workflow for clinician review of lesion annotation:** develop a generic workflow to help clinician find missed or wrongly identified lesions.
4. **Heuristic method for the automatic revision of lesion annotation:** develop a generic method to automatically correct series of lesion annotations.
5. **Experimental validation:** apply the lesion matching and the correction workflow methods on three organs dataset, comprising manual and computed lesion annotation.

1.4. Methods Overview

Lesion change analysis: We developed a new lesion matching method for lesion changes analysis in longitudinal medical scans by analyzing consecutive and non-consecutive scans that meet the above requirements. The method inputs the lesion segmentations in the scans (performed manually or obtained automatically) and consists of two steps:

1. *Registration*: a method for automatic aligning the lesion segmentations into the same coordinate system based on both original scans and organ ROI exact segmentations.
2. *Lesion matching*: a method for automatic matching corresponding lesions in the scans based on their spatial location and relative proximity, outputting a lesion changes graph.

Workflow for lesion detection review: We developed a lesion detection correction workflow consisting in showing the clinician suspected areas of the scans and the lesion annotations, identified by clinically unlikely lesion evolutionary patterns of the lesion changes graph. We showed also how this workflow can be used as an automatic heuristic tool to correct lesion annotations.

1.5. Thesis Novelty

The main novel contributions of this thesis are:

1. **Lesion changes graph formalization:** formalization of lesion changes analysis in longitudinal scans in a graph-theoretic framework, where lesions are vertices, matchings are edges and patterns of lesions changes over time are connected components.
2. **Lesion change analysis method:** new classification of seven changes of individual lesions and five patterns of lesion changes based on the lesion changes graph and its connected components and a generic, model-based method for lesion changes analysis that includes the simultaneous matching of lesions in three or more scans.
3. **Workflow for clinician review of lesion annotation:** a new workflow for the analysis and review of lesions and lesion changes in consecutive longitudinal scans.
4. **Heuristic method for the automatic revision of lesion annotation**: automatization of the workflow for clinician review of lesion annotation.
5. **Experimental validation:** experimental results on lesion changes analysis, lesion detection review, and automatic lesion detection revision of metastatic lesions in 67 longitudinal patient studies (262 scans, 2,295 lesions) in lung and liver CT scans and brain MRI scans that demonstrate the performance of the methods.

1.6. Thesis Organization

This thesis is organized as follows. Chapter 2 presents an overview of the state-of-the-art literature of lesion matching methods and lesion changes analysis. Chapter 3 describes the new method and workflow in detail. Chapter 4 describes the evaluation methodology and the experimental results. Chapter 5 concludes with a summary and future research directions.

Chapter 2

Literature Review

Most existing medical image analysis methods focus on individual lesion detection and segmentation in individual scans. Few address lesion matching, identification and classification of lesion changes in longitudinal patient studies, and detection of missed and wrongly identified lesions. Existing approaches for lesion matching in medical images include similarity-based, (Section 2.1) and overlap-based methods (Section 2.2). A review of computer vision methods for multi-object tracking on 2D images and videos is presented in Section 2.3 Section 2.4 presents hindsight bias, a phenomenon that we use for lesion annotations review.

2.1. Similarity-based methods

Similarity-based methods pair lesions based on intensity and/or shape features. Yan et al. (2018) describes a method that maps lesions to an embedding space and uses the distances between the lesions as a similarity measure. Rafael-Palou et al. (2021) describes a method that matches nodules in lung CT scans with a Siamese Neural Network (SNN) that detects similar embedding feature vectors for pairs of candidate lesions with a matching rate of 88.8% (38 patients, 76 nodules). Cai et al. (2021) presents a method that performs affine scan registration and uses a SNN with a Gaussian heatmap at the lesion centroid in one scan to find the matching lesion in the other scan; its lesion matching rate is 88.7% (480 lesion pairs). Tang et al. (2022) describes an attention-based transformer network based on the lesion center points with a lesion matching rate of 95.3%. These methods assume that organs and lesions undergo minor changes and cannot handle unmatched lesions and complex lesion changes.

2.2. Overlap-based methods

Overlap-based methods match lesions in pairs of registered scans according to their location. Beyer et al. (2004) describes RAM, a semi-automatic method to match manually annotated lung nodules in follow-up CT scans with an accuracy of 86.3% (11 patients, 190 nodules). Beigelman-Aubry et al. (2007), Lee et al. (2007), Tao et al. (2009) and Koo et al. (2012) describe similar methods with matching accuracies of 66.7-92.7% for pulmonary nodules. Hering et al. (2021) presents a method for follow-up metastatic melanoma lesion detection based on registration. None of these methods handles split and merged lesions, nor can detect new lesions in the follow-up

scan. Santoro-Fernandes et al. (2021) describes a cluster-based method that pairs lesions with an accuracy of 98% (32 patients, 736 lesions) in PET/CT scans.

Recent works describe graph-based methods for lesion matching and analysis of lesion changes in volumetric medical images. Kuckertz et al. (2021; 2022) presents a graph-based method for visualizing lesion evolution; it does not provide an analysis of lesion changes and does not evaluate the lesion matching accuracy. Szeskin et al. (2023) and Rochman et al. (2023) describe graph-based methods for lesion changes analysis in pair of scans that includes the classification of individual lesion changes with a reported accuracy of 96% on 57 pairs of patient scans; since the methods analyze pairs of scans, they cannot detect missed and wrongly classified lone lesions in non-consecutive scans. Di Veroli et al. (2023) extends the methods to studies with ≥3 scans. However, it does not provide a method for clinician review and automatic revision of lesion annotations and was only validated on lung and liver CT scans.

2.3. Multi-object tracking

Lesion matching and lesion tracking can be viewed as a multi-object tracking task, for which numerous methods have been developed for optical images and videos (Bolme et al., 2010; Li et al., 2019; Teed et al., 2020; Li et al., 2020; Lee et al., 2021) and for cell tracking in fluorescence microscopy (Padfield et al., 2011). However, these methods are unsuited to analysis of lesion changes in longitudinal patient studies with volumetric scans because they are designed for registered 2D images with very close time intervals and minor object appearance differences. Moreover, they do not handle objects that merge and split.

2.4. Hindsight bias

Providing information to radiologists and/or bringing their attention to suspicious findings is helpful in the revision of lesion detection and in the analysis of lesion changes. This is a type of hindsight bias, which is well known in radiological practice and research (Berlin, 2000; Erly et al., 2010; Hovda et al., 2023). Schill et al. (2023) demonstrated that hindsight bias is both a decision-level and a perceptual phenomenon. Bringing the attention of radiologists to suspicious lesions enables them to detect findings with higher confidence. We adopt this phenomenon to the revision of missed and wrongly identified lesions in longitudinal scans. To the best of our knowledge, this thesis is the first to present a clinician workflow for lesion detection review and an automatic method for lesion detection revision based on it.

Chapter 3

Methods

We present a new method and a new workflow for the analysis and review of lesions and lesion changes in consecutive longitudinal volumetric scans of a patient. The generic model-based method consists of lesion matching, classification of changes in individual lesions, and detection of patterns of lesions changes. The workflow guides clinicians in the detection of missed and wrongly identified lesions in manual and computed lesion annotations using the analysis of lesion changes. It also serves as a heuristic method for the automatic revision of lesion annotations in longitudinal scans.

The tasks are formalized in a graph-theoretic framework in which vertices represent lesions, edges represent lesion matchings (Fig. 2), vertex labels represent classes of changes in individual lesions, and classification of connected components represents patterns of lesions changes (Fig. 3). Lesion matchings are computed with an overlap-based lesion pairing method after establishing a common reference frame by deformable registration of the scans and organ segmentations. The classification of changes in individual lesions and of patterns of lesions changes are computed from the properties of the graph and its connected components. We define seven classes of changes in individual lesions and five patterns of lesions changes that fully describe the evolution of lesions over time. The workflow indicates how the labels of vertices (lone lesion) and the non-consecutive edges are used to detect missed lesions (false negatives) and wrongly identified lesions (false positives). We describe the problem formalization, the method for analysis of lesion changes, the workflow for lesion detection review, and the heuristic method for lesion detection revision next.

The graph-theoretic formalization is described in Section 3.1, the lesion changes analysis is presented Section 3.2. Section 3.3 propose the details of the lesion matching algorithm and Section 3.4 explains the methods for the classification of changes in individual lesions and of patterns of lesions changes. In Section 3.5 we present the workflow for the clinician review of lesion annotations and in Section 3.6 we propose an automated version of the workflow that can be used as a heuristic method.

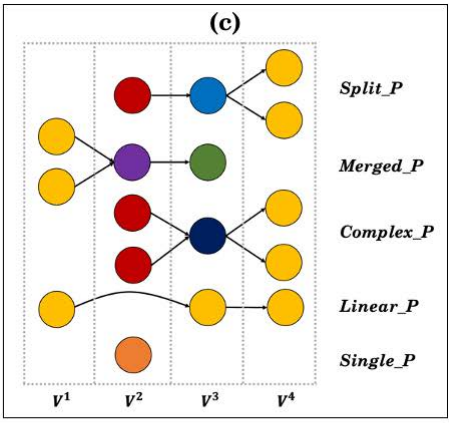
3.1. Problem formalization

The graph-theoretic framework is as follows. Let $S =< S^1, \dots, S^N >$ be a series of $N \geq 3$ consecutive patient scans acquired at times t_i, $1 \leq i \leq N$. Let $G = (V, E)$ be a directed acyclic graph where $V = \{V^i\}, 1 \leq i \leq N$ and $V^i = \{v_1^i, v_2^i, \dots, v_{n_i}^i\}$ is a set of vertices v_j^i corresponding to the lesions associated with the lesion segmentation masks $L^i = \{l_1^i, l_2^i, \dots, l_{n_i}^i\}$, where $n_i \geq 0$ is the number of lesions in scan S^i at time t_i. By definition, any two lesions $v_j^i, v_l^i, j \neq l$ in L^i are disjoint in their voxels. Let $E = \{e_{j,l}^{i,k} = (v_j^i, v_l^k) \mid v_j^i \in V^i, v_l^k \in V^k, \ 1 \leq i < k \leq N\}$ be a set of forward-directed edges connecting vertices in V^i to vertices in V^k. Edge $e_{j,l}^{i,k}$ indicates that the lesions corresponding to vertices v_j^i, v_l^k match, i.e., are the same lesion. This occurs when a lesion appears in scans S^i and S^k in the same anatomical location. Edges between consecutive scans S^i, S^{i+1} are called consecutive edges; edges of non-consecutive scans, $S^i, S^k, i + 1 < k$, are called non-consecutive edges. The in- and out-degree of a vertex v_j^i, $d_{in}(v_j^i)$ and $d_{out}(v_j^i)$, are the number of incoming and outcoming edges, respectively.

Let $CC = \{cc_m\}_{m=1}^M$ be the set of connected components of the undirected graph version of G, where M is the number of connected components and $cc_m = (V_m, E_m)$ is a sub-graph of G, $V_m \subseteq V$, $E_m \subseteq E$. By definition, for each $1 \leq m \leq M$, the sets V_m, E_m are mutually disjoint and their unions are V, E, respectively. In a connected component cc_m, there is an undirected path between any two vertices v_j^i, v_l^k consisting of a sequence of undirected edges in E_m. In this setup, connected components correspond to matched lesions and their pattern of evolution over time.

Changes in individual lesions. We define seven mutually exclusive classes: ***Lone***, ***New***, ***Disappeared***, ***Persistent***, ***Merged***, ***Split*** and ***Complex*** (Fig. 3a). The classes for lesion v_j^i in scan S^i are defined by the number of their predecessors and successors, i.e., the number of lesions matched to v_j^i in previous scans $S^k, 1 \leq k < i$ and in subsequent scans $S^l, i < l \leq N$. The number of predecessors and successors of a lesion are defined by the in- and out-degrees of the vertex associated to the lesion, respectively. For consistency, we assume that each lesion in the first scan S^1 had a predecessor and that each lesion in the current scan S^N has successors. Thus, we set $d_{in}(v_j^1) = 1, \ d_{out}(v_j^N) = 1$.

(a)		Changes in individual lesions						
Matched lesions	**Vertex degree**	*Lone*	*New*	*Disappeared*	*Persistent*	*Merged*	*Split*	*Complex*
Num. of predecessors of l_j^i	$d_{in}(v_j^i)$	0	0	1	1	≥2	0,1	≥2
Num. of successors of l_j^i	$d_{out}(v_j^i)$	0	1	0	1	0,1	≥2	≥2



(b)	Patterns of lesions changes				
Conditions	*Single_P*	*Linear_P*	*Merged_P*	*Split_P*	*Complex_P*
$\|\|cc_m\|\|$	1	>1	>2	>2	>3
Changes in individual lesions	*Lone* or *New* or *Disappeared*	*New* or/and *Disappeared* or/and *Persistent*	At least one *Merged*, no *Split*	At least one *Split*, no *Merged*	At least one *Complex* or/and at least one *Split* and one *Merged*

Fig. 3. (a) Classes of changes in individual lesion l_j^i associated with vertex v_j^i based on its in- and out-degree $d_{in}(v_j^i)$, $d_{out}(v_j^i)$; (b) Patterns of lesions changes defined by the size of the connected component, $\|cc_m\|$, and by the classes of changes of vertices in individual lesions; (c) an example of a lesion changes graph of four consecutive scans: circles are vertices; arrows are consecutive (straight) and non-consecutive (arched) edges; circle colors correspond to classes of changes in an individual lesion; patterns of lesions changes are listed on the right.

Patterns of lesions changes. We define five patterns based on the properties of the connected components cc_m of G and on the type of changes in individual lesion (Fig. 3b): ***Single_P***, a connected component $cc_m = \{v_j^i\}$ consisting of a single lesion; ***Linear_P***, a connected component consisting of a simple sequence of vertices; ***Merged_P***, a connected component consisting of a sequence of vertices in which at least one of them is labeled ***Merged***; ***Split_P***, a connected component consisting of a sequence of vertices in which at least one is labeled ***Split***, and ***Complex_P***, all other connected components.

Fig. 3c illustrates an example of changes in individual lesions and patterns of lesions changes.

3.2. Method for analysis of lesion changes

The inputs to method for the analysis of lesion changes are: 1) the longitudinal scans of the patient; 2) the organ segmentation in each scan; 3) the lesion segmentations in each scan. The

outputs are the lesion matchings, the classification of changes of individual lesions, and the patterns of lesions changes. The method consists of three steps: 1) pairwise deformable registration between each pair of scans and organs segmentations; in each pair, the first scan (moving) is registered to the second scan (fixed); 2) lesion matching computation; 3) detection of changes in individual lesions and patterns of lesions changes from the lesion changes graph properties and from the analysis of its connected components.

The input organ and lesion segmentations in each scan be obtained from manual annotations or from organ- and lesion-specific models, e.g., Szeskin et al. (2023). The lesion changes graph $G = (V, E)$ is initialized by creating a vertex for each lesion in each scan; the edges are computed in the lesion matching step. The pairwise deformable registration of pairs of scans and organ segmentations can be performed with a variety of methods, e.g., Szeskin et al. (2023). We describe lesion matching computation and detection of changes in individual lesions and patterns of lesions changes next.

3.3. Lesion matching computation

Lesion matchings are determined by the location and relative proximity of the lesions in two registered scans. The lesion matching rule is ***lesion voxel overlap***: when the lesion segmentation voxels l_j^i, l_l^k of vertices v_j^i, v_l^k overlap, $1 \leq i < k \leq N$, they are matched and the edge $e_{j,l}^{i,k} = (v_j^i, v_l^k)$ is added to E. Lesion matchings are computed first on consecutive pairs of scans, (S^i, S^{i+1}), $1 \leq i < N$ and then on non-consecutive pairs of scans (S^i, S^k), $1 \leq i + 1 < k \leq N$.

Consecutive lesion matching on scans (S^i, S^{i+1}) is performed with an iterative greedy strategy whose aim is to compensate for registration errors (Rochman et al., 2023, Di Veroli et al., 2023). The inputs are the lesion annotation $\{L^i\}_{i=1}^N$ of a patient study with N scans, and the predefined values of parameters r (number of iterations), d (dilation extension) and p (percentage of overlap). The method works as follows: 1) the lesion segmentations in L^i and L^{i+1} are dilated by d voxels; 2) for all pairs of lesions (v_j^i, v_l^{i+1}), the intersection % of their corresponding lesion segmentations (l_j^i, l_l^{i+1}) is computed as max $(|l_j^i \cap l_l^{i+1}|/|l_j^i|, |l_j^i \cap l_l^{i+1}| / |l_l^{i+1}|)$; 3) if the % intersection is $\geq p$, then the edge $e_{j,l}^{i,i+1} = (v_j^i, v_l^{i+1})$ is added to the set E_c (initially empty); 4) the lesion segmentations l_j^i, l_l^{i+1} are removed from L^i, L^{i+1}, respectively. Steps 1-4 are repeated r times. The

	Lesion matching in consecutive scans
1	**Input**: predefined values for parameters r, d, p; lesion annotations $\{L^i\}_{i=1}^N$
2	Initialization: $E_c = \emptyset, V = \{v_j^i : l_j^i \in L^i, i = 1, \dots, N\}$
3	**For all** consecutive pairs $i = 1, \dots, N-1$ **do**
4	**Repeat** r times
5	Dilate by $\boldsymbol{d}$ voxels the lesion segmentations L^i and L^{i+1}
6	**For all** vertices (v_j^i, v_l^{i+1}) **do**
7	Compute the intersection percentage of their corresponding lesion segmentations (l_j^i, l_l^{i+1}) as $\max(\lvert l_j^i \cap l_l^{i+1}\rvert / \lvert l_j^i\rvert, \lvert l_j^i \cap l_l^{i+1}\rvert / \lvert l_l^{i+1}\rvert)$.
8	**If** the lesion segmentation intersection percentage is $\geq p$ **then**
9	Add the consecutive edge $e_{j,l}^{i,i+1} = (v_j^i, v_l^{i+1})$ to E_c.
10	Remove the lesion segmentations $l_j^i,\ l_l^{i+1}$ from L^i, L^{i+1}, respectively
11	**If** L^i or L^{i+1} are empty, stop
12	**Return** $G_C = (V, E_c)$

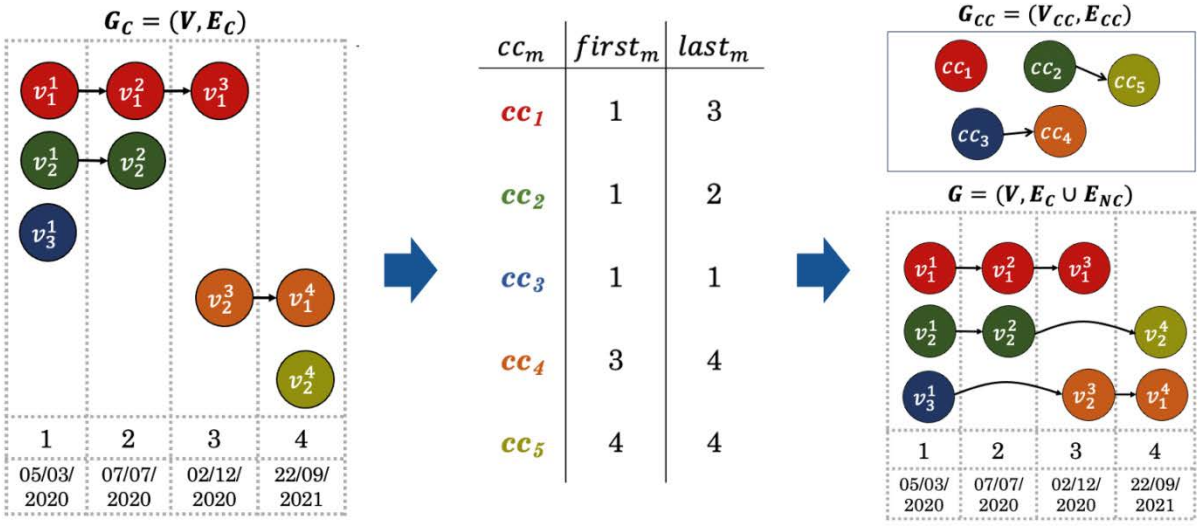
Table 1: Pseudocode for lesion matching in consecutive scans.

values of d, p, r are pre-defined empirically. The output is the graph $G_C = (V, E_C)$, where is E_C the set of consecutive edges $e_{j,l}^{i,i+1}$. Table 1 lists the pseudo-code.

Non-consecutive lesion matching is performed by matching lesions between all ordered pair of non-consecutive scans $(S^i, S^k), 1 \leq i+1 < k \leq N$ and is based on the connected component properties of G_C. First, matches between two candidate lesions are computed. Two lesions (l_j^i, l_l^k) are candidates for a match if their connected components appear in non-consecutive scans and the time elapsed between their appearance is ≤ 1.5 years. Next, deformable registration maps for the scans (S^i, S^k) of the candidate lesions are computed and applied to the lesion pairs. Finally, the registered candidate matching lesions are matched with the consecutive lesion matching method described above. The result is a set of non-consecutive edges E_{NC}. The output of the matching method is the lesion changes graph $G = (V, E), E = E_C \cup E_{NC}$. Table 2 lists the pseudo-code. Fig. 4 illustrates non-consecutive lesion matching.

	Lesion matching in non-consecutive scans
1	**Initialization**: $E_{NC} = \emptyset$ be an empty set of non-consecutive edges.
2	Compute $CC = \{cc_m\}$, the set of connected components of G_C, by Depth First Search (DFS).
3	**For each** $cc_m = (V_m, E_m) \in CC$ compute $first_m$ and $last_m$, the first and last time-indices of the scans in which cc_m's lesions appear, $first_m = \text{argmin}_i\{v_j^i \in V_m\}$; $last_m = \text{argmax}_i\{v_j^i \in V_m\}$.
4	Create $G_{CC} = (V_{cc}, E_{cc})$, the connected component directed acyclic graph, initially empty:
5	**For all** $cc_m \in CC$, add a vertex to V_{CC}
6	**For all** $cc_m \in CC, cc_n \in CC$ **do**
7	**If** $(last_m + 1 < first_n)$ and $(t^{first_n} - t^{last_m} < 1.5 \text{ years})$ The lesions in the two connected cc_m and cc_n appear in non-consecutive scans and the time difference between them is less than 1.5 years
8	**Then** add the edge (cc_m, cc_n) to E_{CC}
9	Create a dictionary D that maps all the ordered pairs of non-consecutive time indices to the candidate lesions for matching. The keys of D are the set of ordered index pairs $\{< i, k > : 1 \leq i + 1 < k \leq N\}$. The values are sets of vertices pairs $\tilde{V}_{i,k}$ computed as follows.
10	**For all** $(cc_m, cc_n) \in E_{CC}$ **do**
10	Add to $\tilde{V}_{last_m, first_n}$ all vertices $v_j^{last_m} \in V_m$ and $v_l^{first_n} \in V_n$
11	**For all** keys $< i, k >$ of D whose value $\tilde{V}_{i,k} \neq \emptyset$ **do**
12	Define candidate lesion masks $\tilde{L}^i = \{l_j^i : v_j^i \in \tilde{V}_{i,k}\}$ and $\tilde{L}^k = \{l_l^k : v_l^k \in \tilde{V}_{i,k}\}$
13	Compute the deformable registration R of scan S^i onto S^k
14	Apply R to $\tilde{L}^i$
15	Compute the set of edges $E_{i,k}$ between $\tilde{L}^i$, $\tilde{L}^k$ (lesion pairings) using the consecutive edges method in Table 1.
16	Add $E_{i,k}$ to E_{NC}
17	**Return** $G = (V, E_C \cup E_{NC})$

Table 2: Pseudocode for lesion matching in non-consecutive scans.



(a) (b) (c)

Fig. 4. Illustration of the non-consecutive matching algorithm in Table 2: (a) Lesion changes graph $G_C = (V, E_C)$ after the computation consecutive edges (Algorithm of Table 1): circles represent vertices, straight arrows represent consecutive edges; vertices with the same color belong to the same connected component. The bottom rows show the time indices and the scan dates; (b) values of $first_m$ and $last_m$ for each connected component cc_m; (c) connected component graph (top): circles represent connected components cc_m, and edges the possible connections between them; final lesion changes graph G (bottom): two non-consecutive edges were added to G_C. The non-consecutive edges were computed by matching the last and first lesions of the cc_m's in G_{CC}.

3.4. Changes in individual lesions and patterns of lesion changes computation

The changes in individual lesions labels are directly computed for each lesion from the resulting graph $G = (V, E)$ with the in- and out-degree of each vertex (Fig. 3a). The patterns of lesions changes are directly computed from the graph by computing the connected components $CC = \{cc_m\}$ of G by standard graph Depth First Search (DFS) and classifying them with the conditions defined in Fig. 3b.

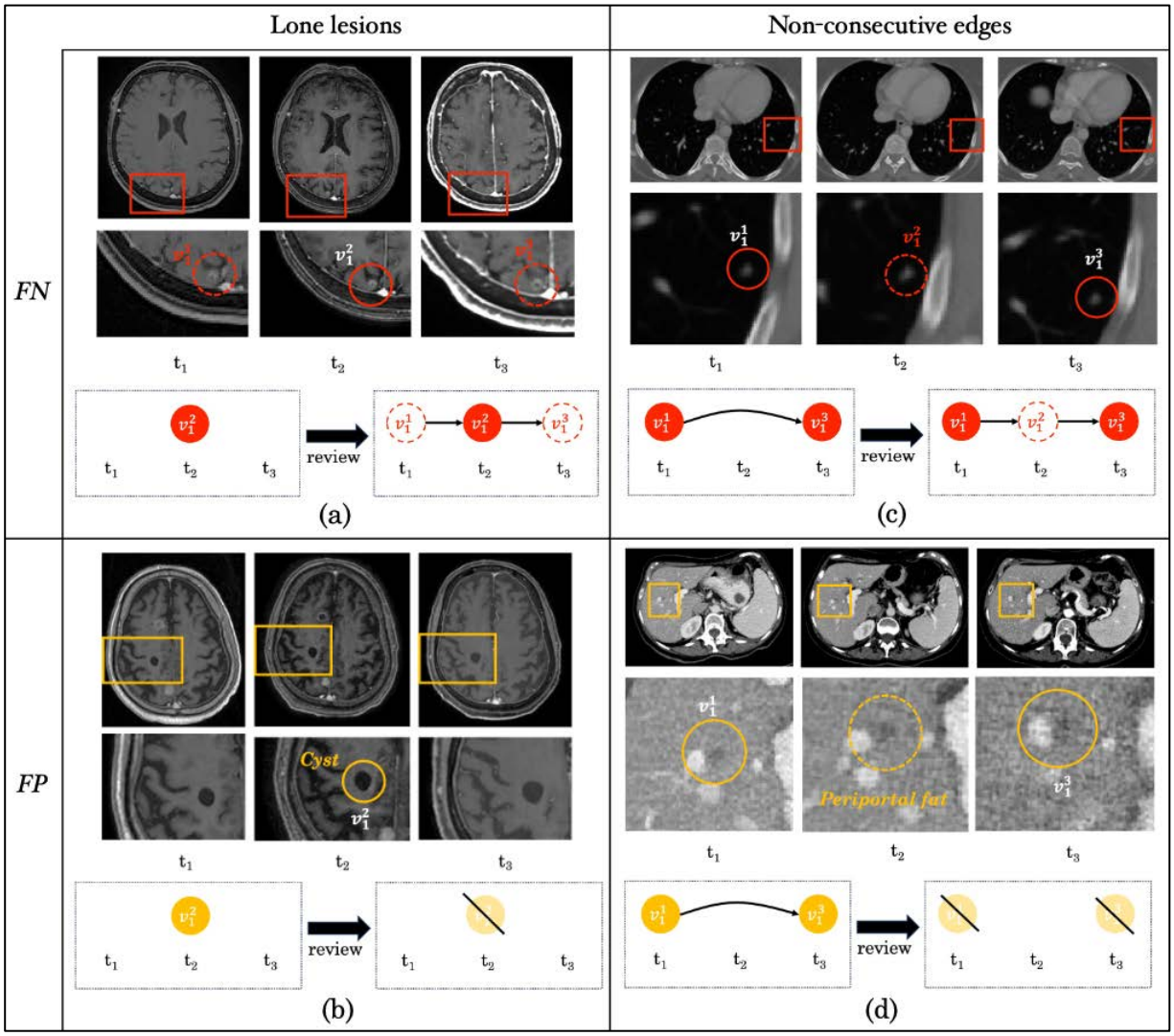


Fig. 5. Illustration of the detection of missed lesions (false negatives FN, top, red) and wrongly identified lesions (false positives FP, bottom, yellow) with lone lesions (left) and non-consecutive edges (right) in four longitudinal scan studies of patients. Each Figure (a)-(d) shows three representative scan slices from consecutive times t_1, t_2, t_3 (top), corresponding enlarged areas (rectangle, bottom) and their lesion changes graph (bottom), before (left) and after (right) the clinician review with the workflow for the review of lesions detection. Solid circles overlaid on the scans indicate the detected lesion area; dotted circles indicate the suspected missed or wrongly identified lesion area. In the graphs, vertex labels indicate the lesion identifier; straight arrows correspond to consecutive edges, curved ones to non-consecutive edges. A cross bar over a vertex indicates the deletion of the lesion and its vertex from the lesion changes graph. Shown are: (a), (b) metastatic brain lesions in T1W-Gad scans; the lesion in (b) is identified in the review as a cyst and not as a brain metastasis; (c) metastatic lung lesions in chest CT scans; (d) metastatic liver lesion in abdominal contrast-enhanced CT scans: the lesion is identified in the review as periportal fat (not as a liver metastasis).

3.5. Workflow for the review of the detection of lesions

The detection of missed (false positive) and wrongly identified lesions (false positives) both in manual ground-truth lesion annotations and in computed lesion segmentations is of great practical importance. Often times, radiologists review the appearance of lesions in the previous and following scans to determine if a faintly seen lesion in one scan also appeared before and/or after. We have identified two scenarios in which the simultaneous analysis of three or more longitudinal scans can help in detecting these lesions which are difficult or cannot be found by examining individual scans or pairs of scans (Fig. 5).

1) Lone lesion: a lesion has been identified in the scan of interest S^i, but not in the previous and the following scans, S^{i-1}, S^{i+1}, $1 < i < N$ (Fig. 5a,b). The lesion may in fact not be a lesion, but a structure that looks like a lesion or a scanning artifact: it is very unlikely that a lesion appeared and disappeared on the same site in-between the time of the scans' acquisitions. Alternatively, the suspected lesion may indeed be a lesion that, due to its size and appearance, shows faintly or not at all on the previous and/or next scans. By examining the region where the suspected lesion appears in all three scans, the clinician may agree that either the lesion is a lone lesion (true positive) and then there may be a matching missed lesion in the previous and/or following scan (false negatives), or that it was wrongly detected in the current scan (false positive). This scenario corresponds to the changes in individual lesion label ***Lone***.

2) Non-consecutive matched lesions: a lesion has been identified and matched in two non-consecutive scans S^i, S^k, $1 \leq i + 1 < k \leq N$, but has not been identified in the in-between scans S^l, $1 \leq i < l < k \leq N$ (Fig. 5c,d). The lesion may in fact have been missed in the in-between scans due to its small size and/or faint appearance: it is very unlikely that a lesion appeared and disappeared in the time between the scans at the same site. Alternatively, the identified and matched lesions may have been wrongly identified in the scans S^i, S^k. By examining the region where the suspected lesion appears in the matched and intermediate scans, the clinician may determine that either the lesion indeed disappeared in the in-between scans (true negative) and then was wrongly detected in the early and/or later scan S^i, S^k (false positive), or that it was missed in the in-between scans (false negative). This scenario corresponds to the non-consecutive lesion matching edges in the lesion changes graph.

We hypothesize that bringing these two suspicious patterns to the attention of the clinicians may help identify hard-to-find missed lesions and discard wrongly identified lesions that may be

difficult or impossible to find by standalone or pairwise examination of the scans. We thus developed a novel workflow for the detection of missed and wrongly identified lesions in longitudinal scans.

Fig. 6 shows the workflow for the review of lesion detection. The inputs are the longitudinal scans of the patient, a list of lone lesions, a list of non-consecutive edges, the lesion changes graph, and the lesion annotations in all scans. For each lone lesion l_j^i, the clinician reviews the lesion in the scan S^i and in the previous and next scans S^{i-1} and S^{i+1}: if the lesion is a false positive, the lesion and its vertex are removed from the lesion changes graph; otherwise, the lone lesion is a true positive and annotated lesions are searched in the same area in previous and next scans. if annotated matching lesions are found, the corresponding consecutive edges are created; otherwise, lesions which were not previously annotated in the previous and next scans are searched: if missed lesions were found, these are false negatives and their corresponding vertices and consecutive edges are added to the lesion changes graph, and the vertices labels are updated accordingly; otherwise, no action is taken. The workflow for non-consecutive edges is similar. Fig. 5 illustrates its use.

3.6. Automatic revision of lesion annotation in longitudinal scans

The workflow for the review of lesion detection can be used as a heuristic method for the automatic detection of missed and wrongly identified lesions in longitudinal scans in which lesions have been manually or automatically annotated. The inputs are: 1) the longitudinal scans of a patient; 2) the lesion annotations in the scan; 3) the analysis of lesion changes from which lists of lone lesions and non-consecutive matching edges are extracted. The outputs are the revised lesion annotations and the revised analysis of lesion changes from which wrongly identified lesions have been removed and to which missing lesions have been added.

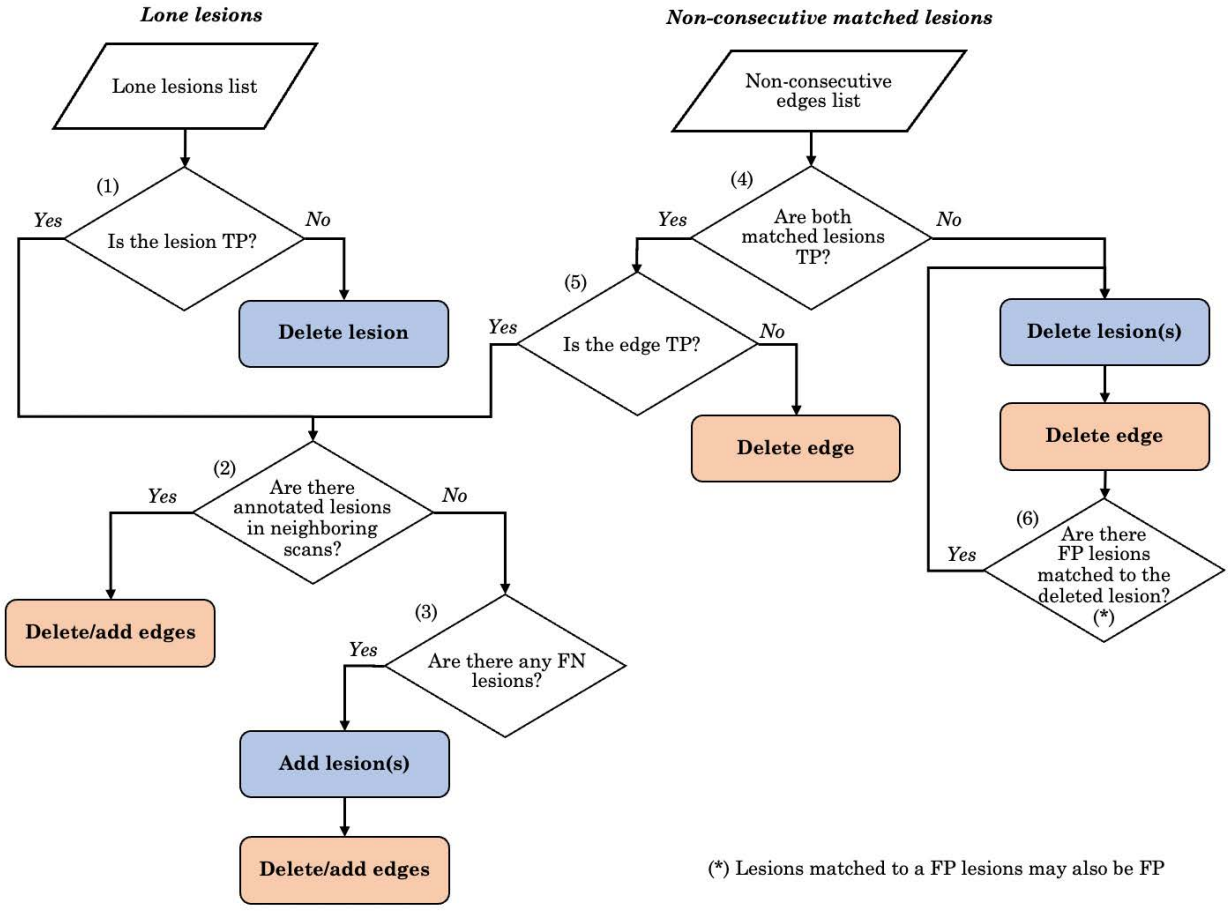


Fig. 6. Workflow for the review of lesion detection. Detection of missed and wrongly identified lesions in longitudinal scans for lone lesions (left) and matched non-consecutive lesions (right). Diamond shapes represent decisions; rounded rectangles represent update operations on the lesion annotations and on the lesion changes graph: add/delete lesion(s), blue; add/delete edge(s), orange. Lesion TP and Edge TP indicate confirmed (true positive) lesion and lesion matches after review by the clinician. Numbers indicate the decisions made by clinicians: lesion detection (1,3,4,6), lesion matching (5), and presence of lesion annotation (2).

The workflow becomes an automatic method by automating the clinical decisions indicated by numbers in Fig. 6. Lesion detection in a scan (decisions 1,3,4,6) is performed with model-based or voxel-based deep learning classification method. To determine if a suspected lesion is true or false positive, the lesion detection module is applied to the suspected lesion region of interest (ROI) in the current scan – an enlarged axis-aligned bounding box containing it. To determine if a suspected lesion is a true or false negative, the lesion detection module is applied to the suspected lesion ROI of the current scans transposed to the previous/next scans, for lone lesions, or to the in-between scans for non-consecutive edges. Lesion matching (decision 5) is performed with the

method described in Section 3.3 with different parameter values for the dilation, intersection %, and number of iterations parameters d, p, r, or with alternative lesion matching methods (Section 2). Presence of a lesion annotation (decision 2) is performed by checking if the lesions appear in the input list of annotated lesions.

Chapter 4

Experimental Results

This chapter describes the experimental studies performed to evaluate our methods. Section 4.1 describes the datasets and Section 4.2 the ground-truth annotations. In Section 4.3 a deep learning for lesion annotation is presented along its results. Section 4.4 lists the evaluation metrics and Section 4.5 the experimental studies results. In Section 4.6 we discuss the results of the experimental studies.

4.1. Datasets

Longitudinal radiological studies of patients with metastatic disease (breast, colon, pancreas) at various cancer stages were retrospectively collected from two Hadassah University Medical Centers (Jerusalem, Israel). The patient studies consist of three or more volumetric scans acquired as part of the routine clinical evaluation of oncology patients with metastatic disease. An IRB waiver for informed consent was obtained. All patient identification and Patient Health Information (PHI) were removed from the data prior to evaluation. Three types of scans were collected: chest CT and contrast-enhanced venous phase abdominal CT (CECT) scans were acquired on Philips CT Brilliance iCT, Canon CT Aquilion Prime SP, and GE CT Optima 660 scanners; brain MRI axial 3D T1W-Gadolinium enhanced scans were acquired on Philips Ingenia, GE Voyager and Siemens Avanto Fit scanners.

Three datasets were created: dataset ***DLUNGS-STUDIES*** consists of 19 patient studies with 83 chest CT scans with a mean 4.4±2.0 scans/patient (maximum 11 scans/patient), a mean time interval between consecutive scans of 125.9±81.3 days, and voxel sizes of 0.6-1.0×0.6-1.0×1.0-3.0mm^3. Dataset ***DLIVER-STUDIES*** consists of 18 patient studies with 77 abdominal CECT scans with a mean 4.3±2.0 scans/patient (maximum 9 scans/patient), a mean time interval between consecutive scans of 109.7±93.5 days, and voxel sizes of 0.6-1.0×0.6-1.0×0.8-5.0mm^3. Dataset ***DBRAIN-STUDIES*** consists of 30 patient studies with 102 brain T1W-Gad MRI scans from with a mean 3.4±0.5 scans/patient (maximum 5 scans/patient), a mean time interval between consecutive scans of 102.1±93.6 days and voxel sizes of 0.2-1.2×0.2-1.2×0.7-4.4mm^3. We refer to them collectively as ***DORGAN-STUDIES.***

4.2. Ground truth lesion annotations and analyses of lesion changes

Ground truth lesion annotations for the ***DLUNGS-STUDIES*** and ***DLIVER-STUDIES*** datasets were manually created by an expert radiologist and for the ***DBRAIN-STUDIES*** dataset by a senior neurosurgeon (both co-authors). This yielded the annotation datasets ***DLUNGS-LESIONS***, ***DLIVER-LESIONS***, ***DBRAIN-LESIONS*** with a total of 1178 lung, 800 liver and 317 brain lesions, with a mean of 14.2±19.1, 10.4±7.9 and 3.1±2.7 lesions/scan (lesions with < 20 voxels were excluded), respectively. We refer to them collectively as ***DORGAN-LESIONS.***

Ground truth analysis datasets of lesion changes ***DLUNGS-ANALYSIS-GT***, ***DLIVER-ANALYSIS-GT***, and ***DBRAIN-ANALYSIS-GT*** consisting of lesion matchings, changes in individual lesions labels and patterns of lesion changes were created with method for the analysis of lesion changes on the corresponding ground-truth lesion annotations datasets ***DORGAN-LESIONS***. The parameter values for lesion matching were $\boldsymbol{d = 1}$ (dilation voxels), $\boldsymbol{p = 10}$% (overlap %), $\boldsymbol{r} = 5,7,10$ (# of iterations) for the lungs, liver and brain lesions, respectively. Then, the radiologist and the neurosurgeon reviewed and corrected the results. Table 3 lists the dataset characteristics. We refer to them collectively as ***DORGAN-ANALYSIS-GT.***

Properties of the lesion changes graph					
Organ	Vertices	Edges	Consecutive edges	Non-consecutive edges	Connected components
LUNGS	1178	663	636 (96%)	27 (4%)	516
LIVER	800	483	458 (95%)	25 (5%)	325
BRAIN	317	156	152 (97%)	4 (3%)	161

(a)

Changes in individual lesions								
Organ	*Lone*	*New*	*Disappeared*	*Persistent*	*Merge*	*Split*	*Complex*	Total
LUNGS	109 9.2%	215 18.2%	51 4.3%	785 66.6%	12 1.0%	6 0.5%	0 0%	1178
LIVER	45 5.6%	185 23.1%	76 9.5%	448 56.0%	27 3.4%	18 2.2%	1 0.1%	800
BRAIN	21 6.6%	58 18.3%	45 14.2%	190 59.9%	2 0.6%	1 0.3%	0 0%	317

(b)

Patterns of lesions changes						
Organ	*Single_P*	*Linear_P*	*Merged_P*	*Split_P*	*Complex_P*	Total
LUNGS	207 40.1%	295 57.2%	9 1.7%	3 0.6%	2 0.4%	516
LIVER	157 48.3%	139 42.8%	13 4.0%	8 2.5%	8 2.5%	325
BRAIN	79 49.1%	79 49.1%	2 1.2%	1 0.6%	0 0%	161

(c)

Table 3. Characteristics of the ground truth datasets in the analysis of lesion changes ***DORGAN-ANALYSIS-GT*** for the lungs, liver and brain patient studies: (a) properties of the lesion changes graph: number of vertices, edges, connected components, consecutive and non-consecutive edges and their ratio (%) of the total number of edges; (b) number of vertices changes in individual lesions labels and (%) of the total number of vertices; (c) patterns of lesions changes labels and the (%) of the total of connected components.

4.3. Computed lesion annotations and analyses of lesion changes

Computed lesions annotations for the ***DORGAN-STUDIES*** datasets were automatically created with deep learning models. Datasets ***DORGAN-TRAIN*** of scans of **different patients** from those in ***DORGAN-STUDIES***, were collected at the same institution and in the same setup as described in Section 4.1. Dataset ***DLUNGS-TRAIN*** consists of 172 chest CT scans from 86 patients; dataset ***DLIVER-TRAIN*** consists of 92 CECT scans from 43 patients; dataset ***DBRAIN-TRAIN*** consists of 456 T1W-Gadolinium enhanced MRI scans from 351 patients. Lesion annotations for each dataset were created in the same manner as for ***DORGAN-LESIONS*** (Section 4.2), yielding a total of 4192 lungs, 2127 liver and 1350 brain lesions with a mean of 24.4±47.5, 23.1±36.0 and 3.0±4.0 lesions/scan (lesions with < 20 voxels were excluded).

Three full resolution 3D nnU-Net models (Isensee et al., 2020), **MORGAN**, were built for voxel-level lesion classification and trained on ***DORGAN-TRAIN*** and their respective lesion annotations (one model for each organ dataset). Lesion-level segmentations were computed with the models by assigning a unique label to each 6-neighbor connected region to the resulting output binary mask. The models were used to compute the lesion annotations of ***DORGAN-STUDIES***, referred to as ***DORGAN-LESIONS-C.*** The method for the analysis of lesion changes was then applied to automatically compute the corresponding datasets from the analysis of lesion changes, ***DORGAN-ANALYSIS-C.***

Table 4 lists ***DORGAN-LESIONS-C*** detection and segmentation performance of the three **MORGAN** models with respect to the manual ground truth lesion annotations ***DORGAN-LESIONS*** for each organ. Note that the training sets ***DORGAN-TRAIN*** are fully disjoint (different patients, scans and lesions) from the test set ***DORGAN-STUDIES.***

Deep learning models performance: lesion detection and segmentation							
Organ	***Num. of scans***	**Lesion Detection**					**Lesion Segmentation**
		TP	*FP*	*FN*	*Precision*	*Recall*	*Mean Dice score per scan*
LUNGS	83	819	150	359	0.85	0.70	0.86 ± 0.11
LIVER	77	576	206	228	0.73	0.72	0.74 ± 0.20
BRAIN	102	281	178	37	0.61	0.88	0.87 ± 0.11

Table 4. Results of the lesion detection and segmentation performed by the **MORGAN** models on the scans of ***DORGAN-STUDIES*** with respect to their manual ground truth lesion annotations ***DORGAN-LESIONS*** for each organ. Column 2 lists the number of scans; columns 3-7 list for the lesion detection: the number of True Positive (*TP*), False Positive (*FP*) and False Negative (*FN*) lesions, and Precision and Recall. Column 8 lists the lesion segmentation Dice score (mean and standard deviation).

4.4. Evaluation metrics

The computed and ground truth analyses of lesion changes were evaluated with three metrics.

1) Changes in individual lesion metric (vertices) is the accuracy of the computed changes in individual lesion labels with respect to the ground-truth labels: when the computed vertex label is identical to the ground-truth, the label is correct, otherwise it is wrong. The classification accuracy is the % of correct labels to the total number of ground-truth labels.

2) Lesion matching metric (edges) is the precision and recall of the computed and ground truth edges in the lesion changes graph: True Positive (TP) when the edge is present in both graphs, False Positive (FP) when it is present in the computed graph but not in the ground truth graph, and FN if it is present in the ground truth graph but not in the computed graph. A computed non-consecutive edge is TP when there is a directed path between its vertices in ground truth graph. Formally, let $G_{comp} = (V_{comp}, E_{comp})$ and $G_{gt} = (V_{gt}, E_{gt})$ be the computed and ground truth lesion changes graphs. Then:

$$Edge\_precision = \frac{\left\|E_{gt} \cap E_{comp}\right\| + \left\|\left\{e = \left(v_j^i, v_l^k\right) \in (E_{comp} \setminus E_{gt}), k > i+1 \,\middle|\, v_j^i \to v_l^k \in V_{gt}\right\}\right\|}{\left\|E_{comp}\right\|}$$

$$Edge\_recall = \frac{\left\|E_{gt} \cap E_{comp}\right\| + \left\|\left\{e = \left(v_j^i, v_l^k\right) \in \left(E_{gt} \setminus E_{comp}\right), k > i+1 \;\middle|\; v_j^i \to v_l^k \in V_{comp}\right\}\right\|}{\left\|E_{gt}\right\|}$$

where "‖. ‖" is the size of the set, "\" is the set difference operator and "$v \to u$" indicates a directed path between vertices v and u.

3) Patterns of lesions changes classification metric (connected components) is the % of the connected components CC_{comp} in the computed lesion changes graph with the same patterns of lesions changes as their counterparts CC_{gt} in the ground-truth lesion changes graph. The pattern comparison is performed between pairs of connected components with the same vertices. All other connected components are counted as wrongly classified.

4.5. Experimental studies

We designed and conducted four experimental studies. Experimental studies 1 and 2 quantify the performance of method for the analysis of lesion changes on the ground truth and on the computed lesions annotations. Experimental studies 3 and 4 quantify the performance of the revision of workflow applied by a clinician and of the automatic lesion detection of the ground-truth and computed lesion annotations.

Study 1: comparison of the results for the method of the analysis of lesion changes applied to the manual lesion annotations ***DORGAN-LESIONS*** to the ground truth ***DORGAN-ANALYSIS-GT.*** Table 5 lists the results.

Study 2: comparison of the results for the method of analysis of the lesion changes applied to the computed lesion annotations ***DORGAN-LESIONS-C, DORGAN-ANALYSIS-C*** to the ground truth ***DORGAN-ANALYSIS-GT.*** Table 6 lists the results.

Study 3: comparison of the results of the lesion detection review workflow used by a clinician on the manual lesion annotations ***DORGAN-LESIONS*** to the ground truth ***DORGAN-ANALYSIS-GT.*** Table 7 lists the results.

Study 4: comparison of the results of the automatic lesion detection review method on the computed lesion annotations ***DORGAN-LESIONS-C*** to the ground truth ***DORGAN-ANALYSIS-GT.*** The automation of clinical decisions was performed using ***DORGAN-LESIONS***. Table 8 lists the results. Table 9 lists the number of wrongly identified lesions (false positives, FP) and missed

ANALYSIS of LESION CHANGES -- MANUAL LESION ANNOTATIONS										
Organ	**Lesion matchings**						**Changes in individual lesions**		**Patterns of lesions changes**	
		TP	*FP*	*FN*	*Precision*	*Recall*				
LUNGS	*Consecutive*	632	10	4	0.98	0.99	*Correct*	1145	*Correct*	486
	Non-consec	26	5	1	0.84	0.96	*GT*	1178	*GT*	516
	All	658	15	5	0.98	0.99	*Accuracy*	0.97	*Accuracy*	0.94
LIVER	*Consecutive*	423	30	35	0.93	0.92	*Correct*	695	*Correct*	259
	Non-consec	18	7	7	0.73	0.72	*GT*	800	*GT*	323
	All	441	38	42	0.92	0.91	*Accuracy*	0.87	*Accuracy*	0.80
BRAIN	*Consecutive*	143	0	9	1.00*	0.94	*Correct*	297	*Correct*	151
	Non-consec	3	0	1	1.00*	0.75	*GT*	317	*GT*	161
	All	146	0	10	1.00*	0.94	*Accuracy*	0.94	*Accuracy*	0.94

* > 0.999

Table 5. Results of the method for the analysis of lesion changes on ground-truth lesion annotations for the lungs, liver, and brain ***DORGAN-LESIONS*** datasets. For lesion matchings, the number of True Positive (*TP*), False Positive (*FP*) and False Negative (*FN*) edges, Precision and Recall are listed for consecutive (*Consecutive*), non-consecutive (*Non-consec*) and all (*All*) edges. For changes in individual lesion classification and for patterns of lesions changes classification, the total number of correct classifications (*Correct)*, ground-truth classes (*GT*) and their ratio (*Accuracy*) are listed.

lesions (false negatives, FN) in suspicious patterns and those associated with other patterns. It quantifies the efficacy of the automatic lesion detection revision method in finding FP and FN lesions.

ANALYSIS of LESION CHANGES -- COMPUTED LESION ANNOTATIONS										
Organ	**Lesion matchings**						**Changes in individual lesions**		**Patterns of lesions changes**	
		TP	*FP*	*FN*	*Precision*	*Recall*				
LUNGS	*Cons*	443	6	1	0.99	1.00*	*Correct*	710	*Correct*	238
	Non-cons	5	10	7	0.87	1.00*	*TP*	813	*TP*	272
	All	448	16	8	0.98	1.00*	*Accuracy*	0.87	*Accuracy*	0.88
LIVER	*Cons*	300	11	13	0.96	0.96	*Correct*	450	*Correct*	122
	Non-cons	13	9	3	0.95	0.92	*TP*	561	*TP*	162
	All	313	20	16	0.96	0.96	*Accuracy*	0.80	*Accuracy*	0.75
BRAIN	*Cons*	126	0	7	1.00*	0.95	*Correct*	242	*Correct*	106
	Non-cons	0	2	2	1.00*	1.00*	*TP*	279	*TP*	128
	All	126	2	9	1.00*	0.95	*Accuracy*	0.87	*Accuracy*	0.83

* > 0.999

Table 6. Results of the method for the analysis of lesion changes on computed lesion annotations for the lungs, liver, and brain ***DORGAN-LESIONS-C*** datasets. The labels are as in Table 5. Note that the accuracy of changes in the individual lesions classification and of the patterns of lesions changes is computed with respect to the number of shared (*TP*) vertices and connected components of the computed and ground-truth lesion changes graph.

LESION DETECTION REVIEW -- MANUAL LESION ANNOTATIONS										
Organ	**Lone lesions (diameter > 5mm)**				**Non-consecutive matching lesions**				**Total**	
	FP	*FN*	*Before review*	*After review*	*FP*	*FN*	*Before review*	*After review*	*FP* (%)	*FN* (%)
LUNGS	1	20	55	36	0	37	27	0	1 (0.0%)*	57 (4.8%)
LIVER	8	20	37	13	10	32	25	0	18 (2.3%)	52 (6.5%)
BRAIN	1	7	15	9	0	4	4	0	1 (0.3%)	11 (3.5%)

Table 7. Results of the lesion detection review workflow used by a clinician on the manual lesion annotations for the lungs, liver, and brain ***DORGAN-LESIONS*** datasets. For lone lesions (diameter > 5mm), the number of wrongly detected lesions (*FP*), missed lesions (*FN*), and the total number of lone lesions before and after review. For non-consecutive matching lesions, the number of wrongly detected lesions (*FP*), missed lesions (*FN*), and total number of non-consecutive edges before and after the review. The last column lists the total number of *FP* and *FN* in both categories identified in the review and their ratio (%) to the total number of lesions in ***DORGAN-LESIONS.***

AUTOMATIC LESION DETECTION REVISION -- COMPUTED LESION ANNOTATIONS										
Organ	**Lone lesions***				**Non-consecutive matching lesions**				**Total**	
	FP	*FN*	*Before revision*	*After revision*	*FP*	*FN*	*Before revision*	*After revision*	*FP* (%)	*FN* (%)
LUNGS	26	21	84	42	23	9	28	7	49 (5.0%)	30 (3.1%)
LIVER	48	14	74	14	10	8	28	12	58 (7.4%)	22 (2.8%)
BRAIN	49	1	68	18	8	2	7	0	57 (12.4%)	3 (0.6%)

* Lesions of all sizes were included to account for segmentation inaccuracy.

Table 8. Results of the automatic lesion detection revision method on the computed lesion annotations for the lungs, liver, and brain ***DORGAN-LESIONS-C*** datasets. The labels are the same as in Table 7.

AUTOMATIC LESION DETECTION REVISION -- COMPUTED LESION ANNOTATIONS					
Organ	**Wrongly identified lesions (FP)**			*Total FP lesions in suspicious patterns / all FP lesions (%)*	**Missed lesions (FN)**
	Lesions in pattern / all lesions in pattern (%)				*Total FN lesions in suspicious patterns / all FN lesions (%)*
	Lesions in suspicious patterns		*Lesions in other patterns*		
	Lone lesion	*Non-consecutive edges*	*All other lesions*		
LUNGS	26/84 (31%)	23/55 (42%)	101/826 (12%)	49/150 (33%)	30/359 (8%)
LIVER	48/74 (65%)	10/51 (20%)	136/717 (19%)	58/206 (28%)	22/228 (10%)
BRAIN	49/68 (72%)	8/14 (57%)	118/377 (31%)	57/178 (32%)	3/37 (8%)

Table 9. Results of the automatic lesion detection revision method on the computed lesion annotations for the lungs, liver, and brain ***DORGAN-LESIONS-C*** datasets for all wrongly identified lesions (false positives, FP) and missed lesions (false negatives, FN). Columns 2-3: ratio of the number of FP lesions in the lone lesion and non-consecutive edges patterns to the total number of lesions in those patterns and %. Column 4: ratio of the number of FP lesions in other patterns to the total number of lesions other patterns and %. Column 5: ratio of the number of FP lesions in the lone lesion and non-consecutive edge patterns to the total number of FP lesions and %. Column 6: ratio of the number of FN lesions in the lone lesion and non-consecutive edge patterns to the total number of FN lesions and %.

4.6. Discussion

Radiological follow-up of oncology patients requires the detection of metastatic lesions and the quantitative analysis of lesion changes in one or more organs in longitudinal volumetric imaging studies. This is a time-consuming task that requires expertise and is subject to observer variability. While recently published methods and commercial software offer partial support for lesion detection, lesion segmentation, lesion matching and tracking, and classification of lesion changes,

none is applicable to all types and patterns of lesions changes, types of lesions, and imaging modalities.

This thesis presents a new method for analysis of lesion changes and a new workflow and heuristic for the review of detection of lesions and lesion changes in longitudinal volumetric scans of a patient. The method for the analysis of lesion changes finds matches between lesions in pairs of scans producing the lesion changes graph, in which lesions are vertices and matches are edges. Using standard graph algorithms, individual lesions (vertices) and lesion patterns (connected components) are labeled according to their changes in time (Fig. 3). Performing the analysis of lesion changes in three or more scans and lesion matching in non-consecutive scans, we can identify clinically unlikely, suspicious graph patterns, namely, lone lesions and non-consecutive edges (Fig. 2). The workflow presented in this thesis guides the clinical review of the lesions to the scan regions associated with the suspicious patterns, that may contain missed or wrongly identified lesions. The heuristic method consists in automatizing clinical decisions during the review (Fig. 5,6). The analysis of lesion changes and the lesion review workflow and heuristic provide the clinician with a comprehensive and succinct set of tools to accurately assess disease status and support clinical decision making.

Our method has several advantages. First, it is generic, does not require training data, and requires only three predefined parameter values for lesion matching. Moreover, it is fully explainable and applicable to multiple organs and types of lesions, on manually or computed lesion annotations, and on both MRI and CT scans. Our workflow is also generic and explainable, applicable to multiple type of lesions. It can be used by a clinician, or it can be automated and used as a heuristic tool. Both our method and workflow leverage all the information contained in the longitudinal scans of a patient.

Studies 1 and 2 show that the method for the analysis of lesion changes achieves state-of-the-art performance both on manual and computed lesion annotations, particularly for the lungs and brain patient studies. The lower performance for the liver patient studies most likely stems from the fact that the lesion matching iterative dilation may match wrong lesions for large lesions (mean lesion sizes: 10.5±53.6 cm^3 liver, 0.4±1.6 cm^3 lungs, 1.6±5.8 cm^3 brain) for multiple, nearby lesions (lesions per scan: 10.4±7.9 liver, 14.2±19.1 lungs 3.1±2.7 in brain) and lesions in conglomerates. However, the improved lesion matching method described in Section 3.3 outperforms the one

described in Di Veroli et al., 2023 on the liver and lung datasets, with fewer false positives and false negative non-consecutive edges.

Studies 3 and 4 show the effectiveness of the workflow for lesion detection review and of the automatic lesion revision heuristic method in finding missed and wrongly identified lesions. In Study 3, the clinician was asked to review his lesion annotations with the workflow. The result shows that non-consecutive edges are indeed a clinically unlikely pattern since all 56 previously detected non-consecutive edges were deleted or substituted by consecutive edges and a lesion detection error was identified (0 non-consecutive edges after the review). Lone lesions, instead, are plausible but not very common (107 before the review, 58 after). The clinician review revealed 120 previously missed lesions (FN) and 20 wrongly identified lesions (FP): the larger number of FN can be explained by the fact that some of these lesions were faintly visible or surmised to be present due to imaging artifacts, e.g., breathing motion, poor contrast agent diffusion, presence of metal artifacts or image inhomogeneity. Our review workflow provided the hindsight to detect them.

Study 4 shows that the automatic lesion detection revision method detects a significant number of missed lesions (55 FN) and wrongly identified lesions (164 FP), that is 8.8% of all FN and 30.7% of all FP lesions all three datasets. Furthermore, the frequency of FP lesions in suspicious patterns and in other patterns constitutes 47.4% of all the lesions in the suspicious patterns. The percentage of FP was only 18.5% of all the lesions not related to suspicious patterns: the FP percentage differences between suspicious and non-suspicious patterns is statistically significant (p-value < 0.01). This shows that the heuristic lesion detection revision can identify hard-to-find detection errors and can optimize the revision of lesion annotations.

Our experimental studies have the following limitations. First, the datasets include scans from only two institutions. However, these included studies performed with scanners from four vendors. Second, patient studies in which the organ underwent significant changes, e.g., patients who underwent liver surgery or had a liver volume difference > 500cc, patients with pleural effusion or lung collapse were excluded. Third, the individual lesion ground truth was established from manual delineation of lesions, which is subject to error and observer variability, especially for small lesions and lesions with fuzzy contours (Joskowicz et al., 2019). However, no time limit was imposed on the expert radiologist for obtaining the optimal delineations.

Chapter 5

Conclusions

This chapter concludes the thesis. Section 5.1 summarizes the methods and the experimental results. Section 5.2 lists the thesis contributions. Section 5.3 presents direction for future research.

5.1. Summary

This thesis presents a new method for the analysis of lesion longitudinal changes and a novel workflow and heuristic method for the revision of lesion annotations that highlights missed and wrongly identified lesions. The tasks were formalized in a graph-theoretic problems: lesions in patient scans are represented by vertices and lesion matchings, i.e., pairing between the appearance of the same lesion in two scans, are represented by edges. In this framework, we defined seven classes of changes of individual lesions based on the lesion's predecessors and successors. Moreso, the connected components of the graph represent patterns of lesions changes over the course of the patient follow-up: we classified these patterns into five classes that exhaustively describe their time evolution.

Lesion matching is performed by pairwise registration, i.e., aligning the scans into the same coordinate system, and by an overlap-based algorithm, applied on consecutive and non-consecutive scans. Lesion matching output is the set of edges of the lesion changes graph. We then computed the classification of the individual lesion changes and of the patterns of lesions changes with standard graph algorithms.

In the lesion detection review workflow, the clinician is called to review areas of the patient scans associated with clinically unlikely patterns of the lesion changes graph. These patterns, namely, lone lesions and non-consecutive edges, may suggest the presence of a missed lesion or of an annotated finding wrongly identified as lesion. The workflow can be executed by clinicians or it can be used as a heuristic method, by automating clinician decision with lesion detection and matching algorithms.

We validated our methods on three datasets of 67 metastatic patient studies comprising lungs and liver CT scans and brain MRI scans (≥3 scans/patient, in total, 262 scans), and their manually annotated lesions (in total, 2295 lesions). For all three datasets, the computed lesion changes analysis yielded a high precision and recall in edge detection (0.92-1.00 and 0.91-0.99) and high

accuracy in the classification of changes in individual lesions (0.87-0.97) and of pattern of lesion changes (0.80-0.94). Similar results were obtained also when we inputted our method computed lesions annotations. The lesion detection review workflow was executed by an expert radiologist and a neurosurgeon on the three datasets and 120 missed and 20 wrongly identified lesions were discovered, while the application of the heuristic method on computed lesion annotations yielded 55 missed and 164 wrongly identified lesions.

5.2. Contributions

The main contributions of this thesis are: 1) the formalization of lesion changes analysis in longitudinal scans in a graph-theoretic framework; 2) a new classification of individual lesions and patterns of lesion changes based on the lesion changes and a generic, model-based method for lesion changes analysis that includes the simultaneous matching of lesions in three or more scans; 3) a new workflow for the analysis and review of lesion detection and lesion changes in consecutive longitudinal scans; 4) a heuristic method for the lesion detection review, based on the workflow; 5) Four experimental studies evaluating the computed lesion changes analysis, lesion detection review, and automatic lesion detection revision of metastatic lesions in 67 longitudinal patient studies.

5.3. Future Work

We consider several directions for the future work. First, improve the registration process by developing a groupwise registration capable of aligning series of 3D scans with large time intervals (~3 months) and with organ deformations: this may shorten the amount of time needed for pairwise registration. Second, the matching algorithm can be sophisticated by adapting the lesion mask dilation to its anatomical location within the organ. Third, given additional information about the patient personal details and medical background the lesion change graph can be used to predict the success of a certain therapy.

Further directions for improving the lesion detection workflow can be exploring other suspicious patterns of the lesion change graph to find missed and wrongly identified lesions. For the heuristic method, specific algorithms may be crafted to switch the clinician decision points.

Finally, our methods may be validated on further datasets from different institutions.

References

1. Beigelman-Aubry, C., Raffy, P., Yang, W., Castellino, R.A., Grenier, P.A. 2007. Computer-aided detection of solid lung nodules on follow-up MDCT screening: evaluation of detection, tracking, and reading time. Am. J. Roentgenology 189(4), 948-955.
2. Berlin, L. 2000. Hindsight bias. Am. J. Roentgenology 175(3).
3. Beyer, F., Wormanns, D., Novak, C., Shen, H., Odry, B.L., Kohl, G., Heindel, W. 2004. Clinical evaluation of a software for automated localization of lung nodules at follow-up CT examinations. Rofo: Fortschritte auf dem Gebiete der Rontgenstrahlen und Nuklearmedizin, 176(6), 829-836.
4. Bolme, D.S., Beveridge, J.R., Draper, B.A., Lui, Y.M. 2010. Visual object tracking using adaptive correlation filters. Proc. IEEE Conf. Comp. Vision & Pattern Recognition, pp. 2544-2550.
5. Cai, J., Tang, Y., Yan, K., Harrison, A.P., Xiao, J., Lin, G., Lu, L. 2021. Deep lesion tracker: Monitoring lesions in 4D longitudinal imaging studies. Proc. IEEE Conf. on Computer Vision and Pattern Recognition, pp. 15159-15169.
6. Di Veroli, B., Lederman, R., Sosna, J., Joskowicz, L. 2023. Graph-theoretic automatic lesion tracking and detection of patterns of lesion changes in longitudinal CT studies. Accepted to MICCAI 2023.
7. Eisenhauer, E.A., Therasse, P., Bogaerts, J. 2009. New response evaluation criteria in solid tumors: revised RECIST guideline version 1.1. European J of Cancer, 45(2), 228-247.
8. Erly, W. K., Tran, M., Dillon, R. C., Krupinski, E. 2010. Impact of Hindsight Bias on Interpretation of Nonenhanced Computed Tomographic Head Scans for Acute Stroke. Journal of Computer Assisted Tomography 34(2):p 229-232.
9. Hering, A., Peisen, F., Amaral, T., Gatidis, S., Eigentler, T., Othman, A., Moltz, J. 2021. Whole-Body Soft-Tissue Lesion Tracking and Segmentation in Longitudinal CT Imaging Studies. Proceedings of the Fourth Conference on Medical Imaging with Deep Learning, PMLR 143, 312-326.
10. Hovda, T., Larsen, M., Romundstad, L., Sahlberg, K. K., Hofvind, S. 2023. Breast cancer missed at screening; hindsight or mistakes? European Journal of Radiology, Volume 165.

11. Isensee, F., Jaeger, P.F., Kohl, S.A., Petersen, J., Maier-Hein, K.H. 2020. nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation, *Nature Methods* 1-9.

12. Joskowicz, L., Cohen, D., Caplan, N., Sosna, J. 2019. Inter-observer variability of manual contour delineation of structures in CT. European Radiology 29(3):1391-1399.

13. Joskowicz, L., Szeskin, A., Rochman, S., Dodi, A., Lederman, R., Fruchtman-Brot, H., Azraq, Y., Sosna, J., Follow-up of liver metastases: a comparison of deep learning and RECIST 1.1. 2023. European Radiology (https://doi.org/10.1007/s00330-023-09926-0).

14. Ko, J.P., Betke, M. 2001. Chest CT: Automated Nodule Detection and Assessment of Change over Time—Preliminary Experience. Radiology 218:267-273.

15. Koo, C.W., Anand, V., Girvin, F., Wickstrom, M.L., Fantauzzi, J.P., Bogoni, L., Babb, J.S., Ko, J.P. 2012. Improved efficiency of CT interpretation using an automated lung nodule matching program. Am. J. Roentgenology 199(1):91-95.

16. Kuckertz, S., Weiler, F., Matusche, B., Lukas, C., Spies, L., et al. 2021. A system for fully automated monitoring of lesion evolution over time in multiple sclerosis. SPIE Medical Imaging.

17. Kuckertz, S., Klein, J., Engel, C., Geisler, B., Krass, S., Heldmann, S. 2022. Fully automated longitudinal tracking and in-depth analysis of the entire tumor burden: unlocking the complexity. SPIE Medical Imaging.

18. Lee, J., Jeong, M., Ko, B.C. 2021. Graph Convolution Neural Network-Based Data Association for Online Multi-Object Tracking. IEEE Access 9:114535-114546.

19. Lee, K.W., Kim, M., Gierada, D.S., Bae, K.T. 2007. Performance of a computer-aided program for automated matching of metastatic pulmonary nodules detected on follow-up chest CT. Am. J. Roentgenology 189(5):1077-1081.

20. Li, B., Wu, W., Wang, Q., Zhang, F., Xing, J., Yan, J.S. 2019. Evolution of Siamese visual tracking with very deep networks. Proc. IEEE Conf. Comp. Vision & Pattern Recognition, pp. 16-20.

21. Li, J., Gao, X., Jiang, T. 2020. Graph Networks for Multiple Object Tracking. 2020 IEEE Winter Conference on Applications of Computer Vision (WACV), pp. 708-717.

22. Moltz, J.H., Schwier, M., Peitgen, H.O. 2009. A general framework for automatic detection of matching lesions in follow-up CT. Proc. IEEE Int. Symp. on Biomedical Imaging, pp. 843-846.

23. Padfield, D., Rittscher, J., Roysam, B. 2011. Coupled minimum-cost flow cell tracking for high-throughput quantitative analysis. Medical Image Analysis 15(4):650-668.

24. Rafael-Palou, X., Aubanell, A., Bonavita, I., Ceresa, M., Piella, G., Ribas, V., Ballester, M.A.G. 2021. Re-identification and growth detection of pulmonary nodules without image registration using 3D siamese neural networks. Medical Image Analysis, 67:101823.

25. Rochman, S., Szeskin, A., Lederman, R., Sosna, J., Joskowicz, L. 2023. Graph-Based automatic detection and classification of lesion changes in pairs of CT studies for oncology follow-up. Int. J. Computer Assisted Radiology and Surgery. Published online, Aug 4, 2023 https://doi.org/10.1007/s11548-023-03000-2

26. Santoro-Fernandes, V., Huff, D., Scarpelli, M.L., Perk, T.G., Albertini, M.R., Perlman, S., Yip, S.S.F., Jeraj, R. 2021. Development and validation of a longitudinal soft-tissue metastatic lesion matching algorithm. Physics in Medicine & Biology 66(15): 155017.

27. Schill, H.M., Gray, S.M. and Brady, T.F., 2023. Visual hindsight bias for abnormal mammograms in radiologists. Journal of Medical Imaging 10(S1), pp.S11910-S11910.

28. Shafiei, A., Bagheri, M., Farhadi, F., Apolo, A.B., Biassou, N.M., Folio, L.R., Jones, E.C., Summers, R.M. 2021. CT evaluation of lymph nodes that merge or split during the course of a clinical trial: limitations of RECIST 1.1. Radiology: Imaging Cancer 3(3).

29. Szeskin, A., Rochman, S., Weis, S., Lederman, R., Sosna, J., Joskowicz, L. 2022. Graph-Based Automatic Lesion Matching and Change Detection in Pairs of CT Studies. Medical Image Analysis 41(6): 1755-1765.

30. Tan, R.T., Ikeuchi, K. 2001. Modeling and tracking non-rigid motion with the extended Kalman filter. Proc. IEEE Conf. Comp. Vision & Pattern Recognition, Vol. 2, pp. 592-599.

31. Teed, Z., Deng, J. Raft. Recurrent all-pairs field transforms for optical flow. 2020. Proc. European Conference on Computer Vision, pp. 402-419.

32. Wang, N., Zeng, G., Ji, X., Zhang, L. 2020. Learning dynamic context-aware correlation filters for visual tracking. Proc. IEEE Conf. Comp. Vision & Pattern Recognition, pp. 4956-4965.

33. Zheng, B., Wang, J., Yang, M.H. 2002. Object tracking and change detection. Proc. IEEE Conf. Comp. Vision & Pattern Recognition, Vol. 2, pp. 414-421.