Graph-Theoretic Automatic Lesion Tracking and Detection of Patterns of Lesion Changes in Longitudinal CT Studies

Beniamin Di Veroli[1], Richard Lederman[2], Jacob Sosna[2], and Leo Joskowicz[1](✉)

[1] School of Computer Science and Engineering, The Hebrew University of Jerusalem, Jerusalem, Israel
beniamin.diveroli@mail.huji.ac.il, josko@cs.huji.ac.il
[2] Department of Radiology, Hadassah University Medical Center, Jerusalem, Israel

Abstract. Radiological follow-up of oncological patients requires the analysis and comparison of multiple unregistered scans acquired every few months. This process is currently partial, time-consuming and subject to variability. We present a new, generic, graph-based method for tracking individual lesion changes and detecting patterns in the evolution of lesions over time. The tasks are formalized as graph-theoretic problems in which lesions are vertices and edges are lesion pairings computed by overlap-based lesion matching. We define seven individual lesion change classes and five lesion change patterns that fully summarize the evolution of lesions over time. They are directly computed from the graph properties and its connected components with graph-based methods. Experimental results on lung (83 CTs from 19 patients) and liver (77 CECTs from 18 patients) datasets with more than two scans per patient yielded an individual lesion change class accuracy of 98% and 85%, and identification of patterns of lesion change with an accuracy of 96% and 76%, respectively. Highlighting unusual lesion labels and lesion change patterns in the graph helps radiologists identify overlooked or faintly visible lesions. Automatic lesion change classification and pattern detection in longitudinal studies may improve the accuracy and efficiency of radiological interpretation and disease status evaluation.

Keywords: longitudinal follow-up · lesion matching · lesion change analysis

1 Introduction

The periodic acquisition and analysis of volumetric CT and MRI scans of oncology patients is essential for the evaluation of the disease status, the selection of the treatment, and the response to treatment. Currently, scans are acquired every 2–12 months according to the patient's characteristics, disease stage, and treatment regime. The scan interpretation consists of identifying lesions (primary tumors, metastases) in the affected

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/978-3-031-43904-9_11.

H. Greenspan et al. (Eds.): MICCAI 2023, LNCS 14224, pp. 106–115, 2023.
https://doi.org/10.1007/978-3-031-43904-9_11

organs and characterizing their changes over time. Lesion changes include changes in the size of existing lesions, the appearance of new lesions, the disappearance of existing lesions, and complex lesion changes, e.g., the formation of conglomerate lesions. As treatments improve and patients live longer, the number of scans in longitudinal studies increases and their interpretation is more challenging and time-consuming.

Radiological follow-up requires the quantitative analysis of lesions and patterns of lesion changes in subsequent scans. It 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, quantification of lesion changes is partial and approximate. The RECIST 1.1 guidelines call for finding new lesions (if any), identifying up to the five largest lesions in each scan in the CT slice where they appear largest, manually measuring their diameters, and comparing their difference [1]. While volumetric measures of individual lesions and of all lesions (tumor burden) have long been established as more accurate and reliable than partial linear measurements, they are not used clinically because they require manual lesion delineation and lesion matching in unregistered scans, which is usually time-consuming and subject to variability [2].

In a previous paper, we presented an automatic pipeline for the detection and quantification of lesion changes in pairs of CT liver scans [3]. This paper describes a graph-based lesion tracking method for the comprehensive analysis of lesion changes and their patterns at the lesion level. The tasks are formalized as graph-theoretic problems (Fig. 1). Complex lesion changes include merged lesions, which occurs when at least two lesions grow and merge into one (possible disease progression), split lesions, which occurs when a lesion shrinks and cleaves into several parts (possible response to treatment) and conglomeration of lesions, which occurs when clusters of lesions coalesce. While some of these lesion changes have been observed [4], they have been poorly studied. Comprehensive quantitative analysis of lesion changes and patterns is of clinical importance, since response to treatment may vary among lesions, so the analysis of a few lesions may not be representative.

The novelties of this paper are: 1) identification and formalization of longitudinal lesion matching and patterns of lesion changes in CT in a graph-theoretic framework; 2) new classification and detection of changes of individual lesions and lesion patterns based on the properties of the lesion changes graph and its connected components; 3) a simultaneous lesion matching method with more than two scans; 4) graph-based methods for the detection of changes in individual lesions and patterns of lesion changes. Experimental results on lung (83 CTs, 19 patients) and liver (77 CECTs, 18 patients) datasets show that our method yields high classification accuracy.

To the best of our knowledge, ours is the first method to perform longitudinal lesion matching and lesion changes pattern detection. Only a few papers address lesion matching in pairs of CT/MRI scans [5–13] – none performs simultaneous matching of all lesions in more than two scans. Also, very few methods [3, 14] handle matching of split/merged lesions. Although many methods exist for object tracking in optical images and videos [15–17], they are unsuited for analyzing lesion changes since they assume many consecutive 2D images where objects have very similar appearance and undergo small changes between images. Overlap-based methods pair two lesions in registered scans when their segmentations overlap, with a reported accuracy of 66–98% [3, 5–11,

18]. These methods assume that organs and lesions undergo minor changes, are very sensitive to registration errors, and cannot handle complex lesion changes. Similarity-based methods pair two lesions with similar features, e.g., intensity, shape, location [13–16] with an 84–96% accuracy on the DeepLesion dataset [14]. They are susceptible to major changes in the lesion appearance and do not handle complex lesion changes. Split-and-merge matching methods are used for cell tracking in fluorescence microscopy [19]. They are limited to 2D images, assume registration between images, and do not handle conglomerate changes.

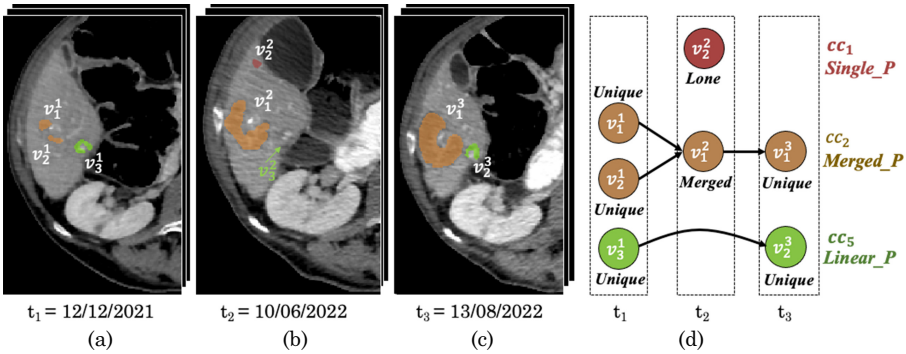


Fig. 1. Longitudinal study of a patient with liver metastases (color overlays): three consecutive (a–c) illustrative slices of unregistered CECT scans acquired at times t_i; matching colors correspond to matching lesions; (d) lesion changes graph: nodes correspond to lesions v_j^i where j is the lesion number at time t_i (dotted rectangle); edges correspond to lesion matches (consecutive straight, non-consecutive curved). The individual lesion changes labels are shown below each node. The graph has three undirected connected components cc_m (red, brown, green), corresponding to three lesion changes patterns, ***Single_P***, ***Merged_P***, ***Linear_P***. Note that lesion v_2^2 labeled as ***Lone*** in t_2 is likely a false positive (not a lesion); lesion v_3^2 (not shown, light green) is likely a false negative (missed lesion) since there is a non-consecutive edge between v_3^1 and v_2^3. (Color figure online)

2 Method

We present a new generic model-based method for the automatic detection and classification of changes in individual lesions and patterns of lesion changes in consecutive CT scans. The tasks are formalized in a graph-theoretic framework in which nodes represent lesions, edges represent lesion matchings, and paths and connected components represent patterns of lesion changes. 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. Changes in individual lesions and patterns of lesion changes are computed from the graph's properties and its connected components. We define seven individual lesion change classes and five lesion change patterns that fully describe the evolution of lesions over time.

The method inputs the scans and the organ and lesion segmentations in each scan. Its outputs are the lesion matchings, the labels of the changes in individual lesions, and the patterns of the lesion changes. The method is a pipeline of four steps: 1) pairwise deformable registration of each prior scan, organ and lesion segmentations, with the most recent (current) scan as in [3]; 2) overlap-based lesion matching; 3) construction of the lesion change graph from the individual lesion segmentations and lesion matches; 4) detection of changes in individual lesions and patterns of lesion changes from the graph properties and from analysis of its connected components.

2.1 Problem Formalization

Let $S = \langle S^1, \ldots, S^N \rangle$ be a series of $N \geq 2$ 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 $V^i = \{v_1^i, v_2^i, \ldots, v_{n_i}^i\}$ is a set of vertices v_j^i corresponding to the lesions associated with the lesion segmentation masks $L^i = \{l_1^i, l_2^i, \ldots, l_{n_i}^i\}$, 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 $E = \left\{e_{j,l}^{i,k} = \left(v_j^i, v_l^k\right) | v_j^i \in V^i, v_l^k \in V^k, 1 \leq i < k \leq N\right\}$ 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 are the same lesion, i.e., that the lesion appears in scans S^i, S^k in the same location. Edges of consecutive scans S^i, S^{i+1} are called consecutive edges; edges of non-consecutive scans, $S^i, S^k, i < k - 1$, 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 (Fig. 1d).

We define **seven** mutually exclusive individual lesion change labels for lesion v_j^i in scan S^i based on the vertex in- and out-degrees (Fig. 2). In the following definitions we refer to the indices: $1 \leq k < i < l \leq N$; 1) ***Lone***: a lesion present in scan S^i and absent in all previous scans S^k and subsequent scans S^l; 2) ***New***: a lesion present in scan S^i and absent in all previous scans S^k; 3) ***Disappeared***: a lesion present in scan S^i and absent in all subsequent scans S^l; 4) ***Unique***: a lesion present in scan S^i and present as a single lesion in a previous scan S^k and/or in a subsequent scan S^l; 5) ***Merged***: a lesion present in scan S^i and present as two or more lesions in a previous scan S^k; 6) ***Split***: a lesion present in scan S^i and present as two or more lesions in a subsequent scan S^l; 7) ***Complex***: a lesion present as two or more lesions in at least one previous scan S^k and at least one subsequent scan S^l. We also define as ***Existing*** a lesion present in scan S^i and present in at least one previous scan S^k and one subsequent scan S^l, $(d_{in}(v_j^i) \geq 1, d_{out}(v_j^i) \geq 1)$. For the first and current scans S^1 and S^N, we set $d_{in}(v_j^1) = 1$, $d_{out}(v_j^N) = 1$, i.e., the

lesion existed before the first scan or remains after the last scan. Thus, lesions in the first (last) scan can only be ***Unique***, ***Disappeared*** or ***Split*** (***Unique***, ***New*** or ***Merged***). Finally, when lesion v_j^i is ***Merged*** and $d_{out}(v_j^i) = 0$, $i < N$, it is also labeled ***Disappeared***; when it is ***Split*** and $d_{in}(v_j^i) = 0$, $i > 1$, it is also labeled ***New***.

We define five patterns of lesion changes based on the properties of the connected components cc_m of G and on the labels of lesion changes: 1) ***Single_P***: a connected component $cc_m = \left\{v_j^i\right\}$ consisting of a single lesion labeled as ***Lone***, ***New***, ***Disappeared***; 2) ***Linear_P***: a connected component consisting of a single earliest vertex v_j^{first} (can be ***New***), a single latest vertex v_l^{last} (can be ***Disappeared***) connected by a sequence. (possibly empty) of ***Unique*** vertices v_j^k, $1 \leq first < k < last \leq N$; ; 3) ***Merged_P***: a connected component whose undirected graph is a tree rooted at a single latest vertex v_l^{last} connected to earlier vertices v_j^k, $1 \leq k < last \leq N$, one or more labeled ***Merged***; 4) ***Split_P***: a connected component whose undirected graph is a tree rooted at a single earliest vertex v_l^{first} connected to later vertices v_j^k, $1 \leq first < k \leq N$, one or more labeled ***Split***; 5) ***Complex_P***: all other connected components. Note that ***Merged_P***, ***Split_P*** and ***Complex_P*** connected components can themselves be subdivided into ***Linear_P***, ***Merge_P***, ***Split_P***, and ***Complex_P*** subcomponents, correspondingly.

Vertex degree	(a) CHANGES IN INDIVIDUAL LESION						
	Lone	*New*	*Disapp-eared*	*Existing*			
				Unique	*Merged*	*Split*	*Complex*
$d_{in}(v_j^i)$	0	0	1	1 (0 first)	≥ 2	0,1	≥ 2
$d_{out}(v_j^i)$	0	1	0	1 (0 last)	0,1	≥ 2	≥ 2
Node pattern							

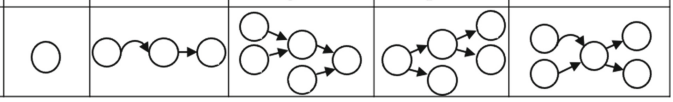
	(b) PATTERNS OF LESION CHANGES				
	Single_P	*Linear_P*	*Merged_P*	*Split_P*	*Complex_P*
Conditions, node labels	*Lone, New, Disappear*	*Single path, at least two nodes*	*Tree, at least one Merge node*	*Tree, at least one Split node*	*All others*
Connected component pattern					

Fig. 2. (a) Individual lesion change classes of a vertex v_j^i defined by its in- and out-degrees $d_{in}(v_j^i)$, $d_{out}(v_j^i)$. Illustrative node pattern: node (circle), edges (arrows); (b) Patterns of lesion changes defined by node labels and connected component properties and illustrative patterns.

The changes in individual lesions and the detection and classification of patterns of lesion changes consist of constructing a graph whose vertices are the corresponding lesion in the scans, computing the graph consecutive and non-consecutive edges that correspond to lesion matchings, computing the connected components of the resulting graph, and assigning an individual lesion change label to each vertex and a lesion change pattern label to each connected component according to the categories above.

2.2 Lesion Matching Computation

Lesion matchings are determined by the location and relative proximity of the lesions in two or more 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 $e_{j,l}^{i,k} = \left(v_j^i, v_l^k\right)$ is added to E. Lesion matchings are computed first on consecutive pairs and then on non-consecutive pairs of scans.

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: 1) the lesion segmentations in L^i and L^{i+1} are isotropically dilated in 3D by d millimeters; 2) for all pairs of lesions $\left(v_j^i, v_l^{i+1}\right)$, compute the intersection % of their corresponding lesion a segmentations (l_j^i, l_l^{i+1}) as $\max(\left|l_j^i \cap l_l^{i+1}\right| / \left|l_j^i\right|, \left|l_j^i \cap l_l^{i+1}\right| / \left|l_l^{i+1}\right|)$; ; 3) if the % intersection is $\geq p$, then edge $e_{j,l}^{i,i+1} = \left(v_j^i, v_l^{i+1}\right)$ is added to E_c; ; 4) remove the lesion segmentations l_j^i, l_l^{i+1} from L^i, L^{i+1}, respectively. Steps 1–4 are repeated r times. This yields the consecutive edges graph $G_C = (V, E_c)$. The values of d, r are pre-defined empirically.

Lesion matching on non-consecutive scans searches for lesion matchings that were not found previously due to missing lesions (unseen or undetected). It is performed by examining the pairs of connected components of G_C and finding possible edges (lesion pairings) between them. Formally, let $CC = \{cc_m\}_{m=1}^{M}$ be the set of undirected connected components of G_C. Let $\tau_m = \left[t_m^{first}, t_m^{last}\right]$ be the time interval between the first and last scans of cc_m, and let $centroid(cc_m)$ be the center of mass of all lesions in cc_m at all times. Let $G_{CC} = (V_{cc}, E_{cc})$ be the graph of connected components of G_C such that each vertex cc_m of V_{cc} corresponds to a connected component cc_m and edges $e_{i,j}^{CC} = \left(cc_i, cc_j\right)$ of E_{CC} satisfy three conditions: 1) the time interval of cc_i is disjoint and precedes by at least one time point that of cc_j, i.e., $t_i^{first} > t_i^{last} + 1$; 2) the connected components are not too far from each other, i.e., the distance between the connected components centroids is smaller than a fixed distance δ, $\left\|centroid(cc_i) - centroid\left(cc_j\right)\right\| \leq \delta$; 3) there is at least one pairwise matching between a lesion in t_i^{last} and a lesion in t_j^{first} computed with the consecutive lesion matching method described above. When the consecutive lesion matching between the lesions in t_i^{last} in cc_i and the lesions in t_j^{first} in cc_j yields a non-empty set of edges, these edges are added as non-consecutive edges to E_c. . Iterating over all non-ordered pairs (cc_i, cc_j) yields the set of consecutive and non-consecutive edges of E.

We illustrate this process with the graph of Fig. 1. First, the consecutive edges (straight lines) of the graph are computed by consecutive lesion matching. This yields a graph with four connected components: $cc_1 = \left\{v_2^2\right\}, cc_2 = \{v_1^1, v_2^1, v_1^2, v_1^3\}, cc_3 = \left\{v_3^1\right\}, cc_4 = \left\{v_2^3\right\}$ corresponding to one ***Merged_P*** and three ***Single_P*** patterns of lesion changes. Connected component cc_2 has no edges to the other connected components since it spans the entire time interval $\tau_1 = [t_1, t_3]$. Connected component cc_1 has no edges to cc_3 and cc_4 since its centroid is too far from them. The connected component pair (cc_3, cc_4) fulfills all three conditions and thus the non-consecutive edge (curved line) $e_{3,2}^{1,3} = (v_3^1, v_2^3)$ is added to E. This results in a new connected component $cc_5 = \{v_3^1, v_2^3\}$ that replaces cc_3 and cc_4 and corresponds to a ***Linear_P*** pattern.

2.3 Classification of Changes in Lesions and in Patterns of Lesion Changes

The changes in individual lesions are directly computed for each lesion from the resulting graph with the in- and out-degree of each vertex (Fig. 2a). The connected components $CC = \{cc_m\}$ of G are computed by graph Depth First Search (DFS). The patterns of lesion changes (Fig. 2b) are computed with path and tree graph algorithms.

The changes in individual lesions, patterns of lesion changes, and lesion changes graph serve as the basis for individual lesion tracking, which consists of following the path from the lesion in the most recent scan backwards to its origins in earlier scans and recording the merged, split and complex lesion changes labels. Summarizing longitudinal studies and queries can also be performed with graph-based algorithms.

3 Experimental Results

We evaluated our method with two studies on retrospectively collected patient datasets that were manually annotated by an expert radiologist.

Dataset: Lung and liver CT studies were retrospectively obtained from two medical centers (Hadassah Univ Hosp Jerusalem Israel) during the routine clinical examination of patients with metastatic disease. Each patient study consists of at least 3 scans.

DLUNG consists of 83 chest CT scans from 19 patients with a mean 4.4 $\pm$ 2.0 scans/patient, a mean time interval between consecutive scans of 125.9 $\pm$ 81.3 days, and voxel sizes of 0.6–1.0 $\times$ 0.6–1.0 $\times$ 1.0–3.0 mm^3. DLIVER consists of 77 abdominal CECT scans from 18 patients with a mean 4.3 $\pm$ 2.0 scans/patient, a mean time interval between consecutive scans of 109.7 $\pm$ 93.5 days, and voxel sizes of 0.6–1.0 $\times$ 0.6–1.0 $\times$ 0.8–5.0 mm^3.

Lesions in both datasets were annotated by an expert radiologist, yielding a total of 1,178 lung and 800 liver lesions, with a mean of 14.2 $\pm$ 19.1 and 10.4 $\pm$ 7.9 lesions/scan (lesions with $<$20 voxels were excluded). Ground-truth lesion matching graphs and lesion changes labeling were produced by running the method on the datasets and then having the radiologist review and correct the resulting node labels and edges.

Study 1: *Lesion changes labeling, lesion matching, evaluation of patterns of lesion changes.* We ran our method on the DLUNGS and DLIVER lesion segmentations. The settings of the parameters were: dilation distance $d = 1$ mm, overlap percentage $p =$ 10%, number of iterations $r = 5$ and 7, and centroid maximum distance $\delta = 17$ and 23 mm for the lungs and liver lesions, respectively.

We compared the computed and ground-truth lesion changes graphs with two metrics: 1) lesion changes classification accuracy, which is the % of correct computed labels from the ground truth labels; 2) lesion matching precision and recall based on the presence/absence of computed vs. ground truth edges. The precision and recall definitions were adapted so that wrong or missed non-consecutive edges are counted as True Positive when there is a path between their vertices in either the ground-truth or the computed graph. Table 1 summarizes the results. The distribution of lesion changes labels for DLUNGS (1,178 lesions) is ***Unique*** 785 (67%), ***New*** 215 (18%), ***Lone*** 109 (9%), ***Disappeared*** 51 (4%), ***Merged*** 12 (1%), ***Split*** 6 (1%), ***Complex*** 0 (0%) with class accuracy $\geq$ 96% for all except ***Split*** (66%). For DLIVER (800 lesions) it is ***Unique*** 450 (56%),

New 185 (23%), ***Lone*** 45 (6%), ***Disappeared*** 77 (10%), ***Merged*** 27 (3%), ***Split*** 18 (2%), ***Complex*** 1 (0.05%) with class accuracy $\geq$ 81% for all except ***Disappeared*** (71%) and ***Split*** (67%).

For the patterns of lesion changes, we compared the computed and ground truth patterns of lesion changes. The accuracy is the % of identical connected components in each category. Table 1 summarizes the results. Note that the ***Split_P***, ***Merged_P*** and ***Complex_P*** patterns jointly account for 3% and 8% of the cases. These patterns are hard to detect manually but their correct classification and tracking are crucial for the proper application of the RECIST 1.1 follow-up protocol [1].

Table 1. (a) Individual lesion change classification (Correctly computed, Ground Truth, Accuracy) and (b) lesion matching (Consecutive, Nonconsecutive, All and Ground Truth, True Positive, False Positive, False Negative, Precision, Recall) results; (c) patterns of lesion changes (Computed, Ground Truth, Accuracy).

Dataset	**(a) Individual lesion change classification**		**(b) Lesions matching Edge detection**						
	Accuracy			*GT*	*TP*	*FP*	*FN*	*Precision*	*Recall*
DLUNGS	*Correct*	1,151	*Cons*	636	628	10	8	0.98	0.99
	GT	1,178	*Noncons*	27	24	2	3	0.96	0.89
	Accuracy	98%	*All*	663	652	12	11	0.98	0.98
DLIVER	*Correct*	679	*Cons*	458	397	39	61	0.91	0.87
	GT	800	*Noncons*	25	17	11	8	0.68	0.80
	Accuracy	85%	*All*	483	414	50	69	0.90	0.86

Dataset	**(c) Patterns of lesion changes**						
		Total	*Single_P*	*Linear_P*	*Merged_P*	*Split_P*	*Complex_P*
DLUNGS	*Computed*	518	210 (41%)	293 (57%)	9 (2%)	3 (1%)	3 (1%)
	GT	516	207 (40%)	295 (57%)	9 (2%)	3 (1%)	2 (0%)
	Accuracy	96%	97%	89%	67%	50%	96%
DLIVER	*Computed*	346	169 (49%)	150 (43%)	11 (3%)	7 (2%)	9 (3%)
	GT	325	157 (48%)	139 (43%)	13 (4%)	8 (2%)	8 (2%)
	Accuracy	76%	85%	76%	54%	0%	25%

Study 2: *Detection of missed lesions in the ground truth.* The expert radiologist was asked to examine non-consecutive edges and lesions labeled as ***Lone*** in the lesion changes graph and determine if lesions were unseen or undetected (actual or presumed false negative) in the skipped or contiguous scans (Fig. 1d). For each non-consecutive edge connecting lesions v_j^i, v_l^k, he analyzed the corresponding region in the skipped scans S^j at $t_j \in]t_i, t_k[$ for possible missed lesions. For the DLUNGS dataset, 25 visible and 5 faintly visible or surmised to be present unmarked lesions were found for 27 non-consecutive edges. For the DLIVER dataset, 20 visible and 21 faintly visible or surmised to be present unmarked lesions were found for 25 non-consecutive edges.

After reviewing the 42 and 37 lesions labeled as ***Lone*** in DLUNGS and DLIVER with > 5mm diameter, the radiologist determined that 1 and 8 of them had been wrongly identified as a cancerous lesion. Moreover, he found that 14 and 16 lesions initially labeled as ***Lone***, had been wrongly classified: for these lesions he found 15 and 21

previously unmarked matching lesions in the next or previous scans. In total, 45 and 62 missing lesions were added to the ground truth DLUNGS and DLIVER datasets, respectively. These hard-to-find ground-truth False Negatives (3.7%, 7.2% of all lesions) may change the radiological interpretation and the disease status. See the Supplemental Material for examples of these scenarios.

4 Conclusion

The use of graph-based methods for lesion tracking and detection of patterns of lesion changes was shown to achieve high accuracy in classifying changes in individual lesion and identifying patterns of lesion changes in liver and lung longitudinal CT studies of patients with metastatic disease. This approach has proven to be useful in detecting missed, faint, and surmised to be present lesions, otherwise hardly detectable by examining the scans separately or in pairs, leveraging the added information provided by evaluating all patient's scans simultaneously using the labels from the lesion changes graph and non-consecutive edges.

References

1. Eisenhauer, E.A., Therasse, P., Bogaerts, J.: New response evaluation criteria in solid tumors: revised RECIST guideline (version 1.1). Eur. J. Cancer **45**(2), 228–247 (2009)
2. Joskowicz, L., Cohen, D., Caplan, N., Sosna, J.: Inter-observer variability of manual contour delineation of structures in CT. Eur. Radiol. **29**(3), 1391–1399 (2019)
3. Szeskin, A., Rochman, S., Weis, S., Lederman, R., Sosna, J., Joskowicz, L.: Liver lesion changes analysis in longitudinal CECT scans by simultaneous deep learning voxel classification with SimU-Net. Med. Image Anal. **83**(1) (2023)
4. Shafiei, A., et al.: CT evaluation of lymph nodes that merge or split during the course of a clinical trial: limitations of RECIST 1.1. Radiol. Imaging Cancer **3**(3) (2021)
5. Beyer, F., et al.: 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 (2004)
6. Lee, K.W., Kim, M., Gierada, D.S., Bae, K.T.: Performance of a computer-aided program for automated matching of metastatic pulmonary nodules detected on follow-up chest CT. Am. J. Roentgenol. **189**(5), 1077–1081 (2007)
7. Koo, C.W., et al.: Improved efficiency of CT interpretation using an automated lung nodule matching program. Am. J. Roentgenol. **199**(1), 91–95 (2012)
8. Tao, C., Gierada, D.S., Zhu, F., Pilgram, T.K., Wang, J.H., Bae, K.T.: Automated matching of pulmonary nodules: evaluation in serial screening chest CT. Am. J. Roentgen. **192**(3), 624–628 (2009)
9. Beigelman-Aubry, C., Raffy, P., Yang, W., Castellino, R.A., Grenier, P.A.: Computer-aided detection of solid lung nodules on follow-up MDCT screening: evaluation of detection, tracking, and reading time. Am. J. Roentgenol. **189**(4), 948–955 (2007)
10. Moltz, J.H., Schwier, M., Peitgen, H.O.: A general framework for automatic detection of matching lesions in follow-up CT. In: Proceedings IEEE International Symposium on Biomedical Imaging, pp. 843–846 (2009)
11. Rafael-Palou, X., et al.: Re-identification and growth detection of pulmonary nodules without image registration using 3D Siamese neural networks. Med. Image Anal. **67**, 101823 (2021)

12. Cai, J., et al.: Deep lesion tracker: monitoring lesions in 4D longitudinal imaging studies. In: Proceedings IEEE Conference on Computer Vision and Pattern Recognition, pp. 15159–15169 (2021)
13. Tang, W., Kang, H., Zhang, H., Yu, P., Arnold, C.W., Zhang, R.: Transformer lesion tracker. arXiv preprint arXiv:2206.06252 (2022)
14. Yan, K., Wang, X., Lu, L., Summers, R.M.: DeepLesion: automated mining of large-scale lesion annotations and universal lesion detection with deep learning. J. Med. Imaging **5**(3), 036501 (2018)
15. Bolme, D.S., Beveridge, J.R., Draper, B.A., Lui, Y.M.: Visual object tracking using adaptive correlation filters. In: Proceedings IEEE Conference Computer Vision & Pattern Recognition, pp. 2544–2550 (2010)
16. Li, B., Wu, W., Wang, Q., Zhang, F., Xing, J., Yan, J.S.: Evolution of Siamese visual tracking with very deep networks. In: Proceedings IEEE Conference Computer Vision & Pattern Recognition, pp. 16–20 (2019)
17. Teed, Z., Deng, J.: RAFT: recurrent all-pairs field transforms for optical flow. In: Proceedings European Conference on Computer Vision, pp. 402–419 (2020)
18. Santoro-Fernandes, V., et al.: Development and validation of a longitudinal soft-tissue metastatic lesion matching algorithm. Phys. Med. Biol. **66**(15), 155017 (2021)
19. Padfield, D., Rittscher, J., Roysam, B.: Coupled minimum-cost flow cell tracking for high-throughput quantitative analysis. Med. Image Anal. **15**(4), 650–668 (2011)