

Novelty Assessment Report

Paper: ProstaTD: Bridging Surgical Triplet from Classification to Fully Supervised Detection

PDF URL: <https://openreview.net/pdf?id=0NkXZ98Bjj>

Venue: ICLR 2026 Conference Submission

Year: 2026

Report Generated: 2026-01-05

Abstract

Surgical triplet detection is a critical task in surgical video analysis, with significant implications for performance assessment and training novice surgeons. However, existing datasets like CholecT50 lack precise spatial bounding box annotations, rendering triplet classification at the image level insufficient for practical applications. The inclusion of bounding box annotations is essential to make this task meaningful, as they provide the spatial context necessary for accurate analysis and improved model generalizability. To address these shortcomings, we introduce ProstaTD, a large-scale, multi-institutional dataset for surgical triplet detection, developed from the technically demanding domain of robot-assisted prostatectomy. ProstaTD offers clinically defined temporal boundaries and high-precision bounding box annotations for each structured triplet activity. The dataset comprises 71,775 video frames and 196,490 annotated triplet instances, collected from 21 surgeries performed across multiple institutions, reflecting a broad range of surgical practices and intraoperative conditions. The annotation process was conducted under rigorous medical supervision and involved more than 60 contributors, including practicing surgeons and medically trained annotators, through multiple iterative phases of labeling and verification. To further facilitate future general-purpose surgical annotation, we developed two tailored labeling tools to improve efficiency and scalability in our annotation workflows. In addition, we created a surgical triplet detection evaluation toolkit that enables standardized and reproducible performance assessment across studies. ProstaTD is the largest and most diverse surgical triplet dataset to date, moving the field from simple classification to full detection with precise spatial and temporal boundaries and thereby providing a robust foundation for fair benchmarking.

Disclaimer

This report is **AI-GENERATED** using Large Language Models and WisPaper (a scholar search engine). It analyzes academic papers' tasks and contributions against retrieved prior work. While this system identifies **POTENTIAL** overlaps and novel directions, **ITS COVERAGE IS NOT EXHAUSTIVE AND JUDGMENTS ARE APPROXIMATE**. These results are intended to assist human reviewers and **SHOULD NOT** be relied upon as a definitive verdict on novelty.

Note that some papers exist in multiple, slightly different versions (e.g., with different titles or URLs). The system may retrieve several versions of the same underlying work. The current automated pipeline does not reliably align or distinguish these cases, so human reviewers will need to disambiguate them manually.

If you have any questions, please contact: mingzhang23@m.fudan.edu.cn

Core Task Landscape

This paper addresses: **Surgical Triplet Detection in Robot-Assisted Prostatectomy Videos**

A total of **8 papers** were analyzed and organized into a taxonomy with **8 categories**.

Taxonomy Overview

The research landscape has been organized into the following main categories:

- **Dataset Development and Annotation Infrastructure**
- **Methodological Approaches for Triplet Recognition and Detection**
- **Clinical Applications and Surgical Practice**

Complete Taxonomy Tree

- Surgical Triplet Detection in Robot-Assisted Prostatectomy Videos Survey Taxonomy
- Dataset Development and Annotation Infrastructure
 - Large-Scale Multi-Institutional Datasets with Bounding Box Annotations ★ (2 papers)
 - [0] ProstaTD: Bridging Surgical Triplet from Classification to Fully Supervised Detection (Anon et al., 2026) [View paper](#)
 - [3] ProstaTD: A Large-scale Multi-source Dataset for Structured Surgical Triplet Detection (Y Chen, 2025) [View paper](#)
 - Holistic Surgical Scene Understanding with Pixel-Wise Recognition (1 papers)
 - [4] Pixel-wise recognition for holistic surgical scene understanding (Nicol ıs Ayobi, 2024) [View paper](#)
- Methodological Approaches for Triplet Recognition and Detection
 - Deep Learning Frameworks for Action Detection (1 papers)
 - [6] A deep learning framework for surgery action detection (Prabu Selvam, 2023) [View paper](#)
 - Triplet Disentanglement and Representation Learning (1 papers)
 - [7] Surgical action triplet recognition via triplet disentanglement (Chen, 2024) [View paper](#)
 - Adversarial Robustness and Uncertainty Handling (1 papers)
 - [1] Guardian: Guarding against uncertainty and adversarial risks in robot-assisted surgeries (Ufaq Khan, 2024) [View paper](#)
- Clinical Applications and Surgical Practice
 - Multi-Console Robotic Surgery Systems and Telesurgery (1 papers)
 - [5] Triple-console robotic telesurgery: first impressions and future impact (T. Rogers, 2024) [View paper](#)
 - Specialized Surgical Techniques and Clinical Case Reports (1 papers)
 - [8] V05-12 SINGLE PORT ROBOT-ASSISTED RADICAL PROSTATECTOMY WITH PERITONEAL FLAP INTERPOSITION IN SETTING OF PRIOR TOTAL   (M Eubank, 2025) [View paper](#)
 - AI and Automation in Endoscopy and Surgery Review (1 papers)
 - [2] Artificial intelligence and automation in endoscopy and surgery (F. Chadebecq, 2023) [View paper](#)

Narrative

Core task: Surgical triplet detection in robot-assisted prostatectomy videos involves identifying and linking three key elements—
instruments, verbs (actions), and targets (anatomical structures)—within surgical video frames. The field's structure, as reflected in the taxonomy, organizes around three main branches: Dataset Development and Annotation Infrastructure, which focuses on creating large-scale, richly annotated video corpora with bounding boxes and triplet labels; Methodological Approaches for Triplet Recognition and

Detection, encompassing algorithmic innovations from deep learning architectures to disentanglement strategies; and Clinical Applications and Surgical Practice, which bridges technical advances to real-world surgical workflows and training. Representative works like ProstaTD Dataset[3] exemplify the dataset development effort by providing multi-institutional annotations, while studies such as Triplet Disentanglement[7] illustrate methodological refinements that decompose the triplet recognition problem into more tractable sub-tasks. Guardian[1] and AI Endoscopy Surgery[2] demonstrate how these technical foundations support clinical decision-making and intraoperative guidance.

A particularly active line of work centers on scaling annotation quality and diversity across institutions, addressing challenges such as inter-annotator variability and the need for pixel-wise versus bounding-box granularity, as seen in Pixel-wise Surgical[4]. Methodological debates revolve around whether to treat triplet detection as a unified end-to-end problem or to disentangle instrument detection, action recognition, and target localization into separate stages. ProstaTD Bridging[0] sits within the Dataset Development and Annotation Infrastructure branch, specifically targeting large-scale multi-institutional datasets with bounding box annotations. Compared to ProstaTD Dataset[3], which established foundational annotation protocols, ProstaTD Bridging[0] appears to extend this infrastructure by addressing cross-institutional harmonization and bridging gaps in annotation consistency. This positions the work as a natural evolution in dataset maturity, complementing methodological advances like those in Deep Learning Action[6] and supporting broader clinical integration efforts exemplified by Triple-console Telesurgery[5] and Single Port Prostatectomy[8].

Related Works in Same Category

The following **1 sibling papers** share the same taxonomy leaf node with the original paper:

1. ProstaTD: A Large-scale Multi-source Dataset for Structured Surgical Triplet Detection

Authors: Y Chen, Z Li, C Xu, AQ Liu, X Xu, et al. (6 authors total) | **Year/Venue:** 2025 | **URL:** [View paper](#)

Abstract

â surgical triplet detection, developed from the technically demanding domain of robot-assisted prostatectomyâ The dataset includes 21 robot-assisted prostatectomy videos, totaling 60,529 â

â Similarity Notice

This paper is highly similar to the original paper; it may be a variant or near-duplicate. Please manually verify.

Contributions Analysis

Overall novelty summary. The paper introduces ProstaTD, a large-scale dataset for surgical triplet detection in robot-assisted prostatectomy, featuring 71,775 frames and 196,490 annotated triplet instances with bounding boxes across 21 multi-institutional surgeries. It resides in the 'Large-Scale Multi-Institutional Datasets with Bounding Box Annotations' leaf, which contains only one sibling paper. This represents a relatively sparse research direction within the broader taxonomy of eight total papers, suggesting the work addresses an emerging need for spatially annotated surgical triplet datasets beyond existing image-level classification resources like CholecT50.

The taxonomy reveals three main branches: Dataset Development, Methodological Approaches, and Clinical Applications. ProstaTD sits within Dataset Development, adjacent to 'Holistic Surgical Scene Understanding with Pixel-Wise Recognition' (one paper) and separate from methodological leaves addressing deep learning frameworks, disentanglement, and adversarial robustness (three papers total). The dataset's multi-institutional scope and bounding box annotations position it as infrastructure enabling the methodological innovations in neighboring branches, while its prostatectomy focus distinguishes it from broader endoscopic surgery datasets. The sparse population of its leaf suggests limited prior work specifically combining large-scale triplet detection with precise spatial annotations across institutions.

Among 29 candidates examined, the analysis identified potential overlap for all three contributions. The core dataset contribution (10 candidates examined, 1 refutable) shows the most novelty, though one prior work appears to provide similar multi-institutional triplet annotations. The annotation tools contribution (9 candidates, 1 refutable) and evaluation toolkit (10 candidates, 2 refutable) face more substantial prior work, with existing open-source labeling frameworks and benchmark protocols identified. These statistics reflect a focused semantic search rather than exhaustive coverage, indicating that within the examined scope, the dataset's scale and domain specificity appear more distinctive than its tooling and evaluation components.

Based on the limited search of 29 candidates, the work's primary novelty appears to lie in its domain-specific scale and multi-institutional scope for prostatectomy triplet detection with spatial annotations. The sparse taxonomy leaf (one sibling) and contribution-level statistics suggest the dataset addresses a genuine gap, though the annotation tools and benchmarking components encounter more established prior work. This assessment reflects top-K semantic matches and may not capture domain-specific precedents outside the search scope.

This paper presents **3 main contributions**, each analyzed against relevant prior work:

Contribution 1: ProstaTD dataset for surgical triplet detection

Description: The authors present ProstaTD, the first large-scale dataset enabling fully supervised surgical triplet detection at the procedure level. It contains 71,775 frames with 196,490 annotated triplet instances from 21 multi-institutional surgeries, featuring precise bounding boxes and clinically defined temporal boundaries for each triplet.

This contribution was assessed against **10 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

1. Estimating surgical urethral length on intraoperative robot-assisted prostatectomy images using artificial intelligence anatomy recognition

URL: [View paper](#)

Brief Assessment

Urethral Length AI[10] focuses on anatomical structure segmentation (prostate, urethra, ligated plexus, catheter) for surgical urethral length estimation, not on surgical triplet detection with instrument-action-target annotations.

2. Pixel-wise recognition for holistic surgical scene understanding

URL: [View paper](#)

Brief Assessment

Pixel-wise Surgical[4] focuses on holistic surgical scene understanding in prostatectomies with instrument segmentation and atomic action detection, not on surgical triplet detection with bounding boxes and triplet boundaries as defined in the original paper.

3. TriQuery: A Query-Based Model for Surgical Triplet Recognition.

URL: [View paper](#)

Brief Assessment

TriQuery[14] full text is not available in the provided context, making comparison impossible. Without access to the candidate paper's content, I cannot assess whether it presents prior work that would refute the novelty of the ProstaTD dataset.

4. Surgical action triplet recognition via triplet disentanglement

URL: [View paper](#)

Brief Assessment

Triplet Disentanglement[7] focuses on a recognition methodology (triplet disentanglement model) rather than dataset construction. While it mentions developing a new dataset from public prostate surgery data, it does not claim to be the first large-scale dataset with precise bounding boxes and temporal boundaries at the procedure level for surgical triplet detection.

5. SAR-RARP50: Segmentation of surgical instrumentation and Action Recognition on Robot-Assisted Radical Prostatectomy Challenge

URL: [View paper](#)

Brief Assessment

SAR-RARP50[9] focuses on surgical tool segmentation and action recognition as separate tasks, not on triplet detection with structured annotations and precise bounding boxes for each triplet instance. The candidate does not provide triplet-level annotations or address the same research problem.

6. Towards holistic surgical scene understanding

URL: [View paper](#)

Brief Assessment

Holistic Surgical Scene[11] focuses on robot-assisted prostatectomy but addresses phase, step, instrument detection, and atomic action recognition tasks, not surgical triplet detection with precise bounding boxes and clinically defined temporal boundaries as in the original paper.

7. ProstaTD: A Large-scale Multi-source Dataset for Structured Surgical Triplet Detection

URL: [View paper](#)

Prior Art Analysis

ProstaTD Dataset[3] demonstrates that a large-scale surgical triplet detection dataset with precise bounding boxes, clinically defined temporal boundaries, and multi-institutional sources already exists. The candidate paper presents ProstaTD with 71,775 frames and 196,490 annotated triplet instances from 21 surgeries across multiple institutions, featuring the same core elements claimed as novel by the original paper: full bounding box annotations, standardized triplet boundaries, and multi-source data collection. Both papers address identical limitations in CholecT50 and propose the same solution architecture.

Evidence

Evidence 1 - **Rationale:** The candidate paper presents an identical dataset with the exact same specifications (71,775 frames, 196,490 instances, 21 surgeries, multi-institutional sources, clinically defined temporal boundaries, and high-precision bounding boxes), demonstrating that this contribution already exists in prior work. - **Original:** we introduce prostatd, a large-scale, multi-institutional dataset for surgical triplet detection, developed from the technically demanding domain of robot-assisted prostatectomy. prostatd offers clinically defined temporal boundaries and high-precision bounding box annotations for each structured tr... - **Candidate:** we introduce prostatd, a large-scale, multi-institutional dataset for surgical triplet detection, developed from the technically demanding domain of robot-assisted prostatectomy. prostatd offers clinically defined temporal boundaries and high-precision bounding box annotations for each structured tr...

Evidence 2 - **Rationale:** Both papers make identical claims about introducing the first large-scale dataset for fully supervised surgical triplet detection at the procedure level, indicating that the candidate paper's work predates or coincides with the original paper's claimed novelty. - **Original:** we introduce a new task with a new dataset for fully supervised surgical triplet detection at the procedure level. to the best of our knowledge, our prostatd is the largest surgical dataset with instance-level annotations. - **Candidate:** we introduce a new task with a new dataset for fully supervised surgical triplet detection at the procedure level. to the best of our knowledge, our prostatd is the largest surgical dataset with instance-level annotations.

8. A deep learning framework for surgery action detection

URL: [View paper](#)

Brief Assessment

Deep Learning Action[6] focuses on action detection in surgical videos but does not provide a large-scale dataset with precise bounding boxes and triplet annotations for prostatectomy. The candidate only mentions 'four complete radical prostatectomy surgery videos,' which is substantially smaller than ProstaTD's 21 surgeries with 71,775 frames and 196,490 annotated triplet instances.

9. A Dataset and Benchmark for Robot-Assisted Radical Prostatectomy With Lymphadenectomy in Surgical Workflow Understanding

URL: [View paper](#)

Brief Assessment

The candidate paper full text is not available (marked as 'n/a'), making it impossible to assess whether it refutes the novelty of ProstaTD as the first large-scale dataset for fully supervised surgical triplet detection at the procedure level.

10. A Dataset for Robot-assisted Radical Prostatectomy with Lymphadenectomy in Surgical Workflow Understanding

URL: [View paper](#)

Brief Assessment

RARP Lymphadenectomy Dataset[13] focuses on phase and step recognition in lymphadenectomy procedures, not surgical triplet detection with instrument-action-target annotations and bounding boxes.

Contribution 2: Open-source annotation tools for surgical triplet labeling

Description: The authors developed two dedicated annotation applications (Triplet-labelme and SurgLabel) specifically designed for surgical triplet annotation. These tools support single-frame triplet editing and high-throughput batch labeling, and will be released as open source to facilitate large-scale annotation across diverse surgical procedures.

This contribution was assessed against **9 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

1. Temset-24k: Densely annotated dataset for indexing multipart endoscopic videos using surgical timeline segmentation

URL: [View paper](#)

Brief Assessment

Temset-24k[16] focuses on timeline segmentation (phase-task-action triplets) for TEMS procedures, not surgical triplet detection (instrument-verb-target). The annotation tools described are Label Studio-based for video timeline annotation, not specialized for surgical triplet bounding box annotation.

2. Grounding Surgical Action Triplets with Instrument Instance Segmentation: A Dataset and Target-Aware Fusion Approach

URL: [View paper](#)

Brief Assessment

Grounding Triplets[27] focuses on creating a dataset for triplet segmentation (linking instrument instance masks with verb-target labels) rather than developing annotation tools for surgical triplet labeling workflows. The paper does not describe dedicated annotation applications for triplet annotation.

3. Surgical Triplet Recognition via Diffusion Model

URL: [View paper](#)

Brief Assessment

Diffusion Triplet[28] focuses on a diffusion-based recognition model for surgical triplets and does not discuss annotation tools or labeling workflows for surgical triplet datasets.

4. Instrument-tissue-guided surgical action triplet detection via textual-temporal trail exploration

URL: [View paper](#)

Brief Assessment

Instrument-tissue Triplet[17] does not describe or mention any annotation tools for surgical triplet labeling. The paper focuses on a detection method (ITG-TRIP) for surgical action triplet detection, not on annotation infrastructure.

5. Frame Selection Methods to Streamline Surgical Video Annotation for Tool Detection Tasks

URL: [View paper](#)

Brief Assessment

Frame Selection Annotation[26] focuses on frame selection methods to streamline surgical video annotation for tool detection tasks in cataract surgery, not on developing annotation tools for surgical triplet labeling workflows. The candidate addresses a different problem (which frames to annotate) rather than providing tools for triplet annotation.

6. Surgical video workflow analysis via visual-language learning

URL: [View paper](#)

Brief Assessment

Visual-language Workflow[24] does not describe annotation tools for surgical triplet labeling. The candidate focuses on a vision-language framework (I2TM) for triplet recognition, not on developing annotation software or tools for labeling workflows.

7. ProstaTD: A Large-scale Multi-source Dataset for Structured Surgical Triplet Detection

URL: [View paper](#)

Prior Art Analysis

ProstaTD Dataset[3] demonstrates that dedicated open-source annotation tools for surgical triplet labeling already exist. The candidate paper presents two annotation applications (Triplet-labelme and SurgLabel) with identical functionality to those claimed as novel by the original paper: single-frame triplet editing and high-throughput batch labeling. Both papers describe these as the first open-source tools specifically designed for surgical triplet annotation.

Evidence

Evidence 1 - **Rationale:** Both papers make identical claims about being the first to release open-source annotation tools specifically tailored for surgical triplet annotation, indicating that the candidate paper's contribution predates the original paper's claimed novelty. - **Original:** we release two open-source annotation tools, which are the first specifically tailored for surgical triplet annotation, along with an open-source evaluation toolkit for benchmarking surgical triplet detection, providing a foundation for surgical triplet analysis across diverse surgical procedures. - **Candidate:** we release two open-source annotation tools, which are the first specifically tailored for surgical triplet annotation, along with an open-source evaluation toolkit for benchmarking surgical triplet detection, providing a foundation for surgical triplet analysis across diverse surgical procedures.

8. Deep-Onto network for surgical workflow and context recognition

URL: [View paper](#)

Brief Assessment

Deep-Onto[25] does not describe annotation tools for surgical triplet labeling. The candidate focuses on surgical workflow recognition using deep learning and ontologies for robot-assisted partial nephrectomy, without addressing triplet annotation infrastructure.

9. Web based Object Annotation Tool using a Triplet-ReID Sorting Approach

URL: [View paper](#)

Brief Assessment

Triplet-ReID Annotation[29] focuses on object re-identification for general datasets using triplet networks to find repeated objects, not surgical triplet annotation (instrument-action-target). The candidate addresses a fundamentally different problem domain (generic object labeling) compared to the surgical workflow annotation tools described in the original paper.

Contribution 3: Evaluation toolkit and benchmark for surgical triplet detection

Description: The authors introduce an evaluation toolkit (ivtdmetrics) tailored for surgical triplet detection benchmarking, supporting metrics such as mAP at various IoU thresholds, precision, recall, and F1-score. They also provide comprehensive benchmarks using state-of-the-art models and propose TDnet as a baseline method.

This contribution was assessed against **10 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

1. Temset-24k: Densely annotated dataset for indexing multipart endoscopic videos using surgical timeline segmentation

URL: [View paper](#)

Brief Assessment

Temset-24k[16] provides benchmarks for timeline segmentation using STALNet, not for surgical triplet detection with bounding boxes. The evaluation focuses on phase/task/action classification accuracy and F1 scores, not mAP/IoU metrics for detection tasks.

2. CurConMix: A Curriculum Contrastive Learning Framework for Enhancing Surgical Action Triplet Recognition

URL: [View paper](#)

Brief Assessment

CurConMix[20] focuses on curriculum contrastive learning for surgical action triplet recognition on cholect45, not on developing evaluation toolkits or benchmarking infrastructure for surgical triplet detection.

3. Instrument-tissue-guided surgical action triplet detection via textual-temporal trail exploration

URL: [View paper](#)

Brief Assessment

Instrument-tissue Triplet[17] presents experimental results on the CholecT50 benchmark but does not introduce a new evaluation toolkit. The paper evaluates their ITG-TRIP method using existing benchmarks rather than proposing new evaluation infrastructure.

4. MT4MTL-KD: A Multi-Teacher Knowledge Distillation Framework for Triplet Recognition

URL: [View paper](#)

Brief Assessment

MT4MTL-KD[23] focuses on a multi-teacher knowledge distillation framework for triplet recognition using the CholecT45 dataset. It does not introduce evaluation toolkits or benchmarking infrastructure for surgical triplet detection.

5. CholecTriplet2021: A benchmark challenge for surgical action triplet recognition

URL: [View paper](#)

Prior Art Analysis

CholecTriplet2021[21] demonstrates that prior work exists on evaluation toolkits and benchmarks for surgical triplet detection. The candidate paper explicitly presents a challenge that assesses state-of-the-art methods for surgical action triplet recognition, providing benchmarking results with mean average precision (mAP) metrics. This directly refutes the novelty claim that the original paper is the first to introduce an evaluation toolkit (ivtdmetrics) and comprehensive benchmarks for surgical triplet detection, as CholecTriplet2021[21] already established such evaluation infrastructure and benchmarking protocols for the surgical triplet detection task.

Evidence

Evidence 1 - **Rationale:** CholecTriplet2021[21] explicitly describes establishing a challenge with assessment infrastructure for surgical action triplet recognition, which directly overlaps with the original paper's claim of creating an evaluation toolkit for surgical triplet detection. - **Original:** we created a surgical triplet detection evaluation toolkit that enables standardized and reproducible performance assessment across studies - **Candidate:** this paper presents cholectriple2021: an endoscopic vision challenge organized at miccai 2021 for the recognition of surgical action triplets in laparoscopic videos. the challenge granted private access to the large-scale cholect50 dataset, which is annotated with action triplet information. in thi...

Evidence 2 - **Rationale:** CholecTriplet2021[21] presents comprehensive benchmarking of multiple methods (23 total) with mAP metrics for surgical action triplet recognition, demonstrating that benchmarking infrastructure and comparative evaluation already existed prior to the original paper's contribution. - **Original:** To enable comprehensive and efficient comparisons, we develop an evaluation toolkit and conduct extensive benchmarks using state-of-the-art models. - **Candidate:** a total of 4 baseline methods from the challenge organizers and 19 new deep learning algorithms from the competing teams are presented to recognize surgical action triplets directly from surgical videos, achieving mean average precision (map) ranging from 4.2% to 38.1%.

Evidence 3 - **Rationale:** CholecTriplet2021[21] demonstrates that methodological comparison and analysis infrastructure for surgical triplet recognition was already established, challenging the claim that the original paper is the first to provide such evaluation toolkit and benchmarking foundation. - **Original:** we release two open-source annotation tools, which are the first specifically tailored for surgical triplet annotation, along with an open-source evaluation toolkit for benchmarking surgical triplet detection, providing a foundation for surgical triplet analysis across diverse surgical procedures. - **Candidate:** this study also analyzes the significance of the results obtained by the presented approaches, performs a thorough methodological comparison between them, in-depth result analysis, and proposes a novel ensemble method for enhanced recognition.

6. Why deep surgical models fail?: Revisiting surgical action triplet recognition through the lens of robustness

URL: [View paper](#)

Brief Assessment

Why Models Fail[22] focuses on robustness analysis and explainability of existing surgical triplet recognition models through adversarial perturbations, not on developing evaluation toolkits or benchmarks for triplet detection. The paper evaluates existing models using standard AP metrics rather than introducing new evaluation infrastructure.

7. ProstaTD: A Large-scale Multi-source Dataset for Structured Surgical Triplet Detection

URL: [View paper](#)

Prior Art Analysis

ProstaTD Dataset[3] demonstrates that an evaluation toolkit and comprehensive benchmark for surgical triplet detection already exists. The candidate paper presents 'ivtdmetrics' toolkit supporting the same metrics (mAP at various IoU thresholds, precision, recall, F1-score) and provides extensive benchmarks using state-of-the-art models, including a baseline method (TDnet), identical to the original paper's claims.

Evidence

Evidence 1 - **Rationale:** Both papers make identical claims about introducing the first benchmark for fully supervised surgical triplet detection with a tailored baseline method, indicating that the candidate paper's contribution predates the original paper's claimed novelty. - **Original:** we introduce the first benchmark for fully supervised surgical triplet detection, providing our tailored method as a baseline for comparison. - **Candidate:** we introduce the first benchmark for fully supervised surgical triplet detection, providing our tailored method as a baseline for comparison.

Evidence 2 - **Rationale:** Both papers present identical benchmark results for TDnet with the exact same performance metrics, demonstrating that the candidate paper's benchmark and baseline method already exist in prior work. - **Original:** our proposed tdnet achieves the best results across all major components. it reaches 36.1% mapivt@0.5 and 33.1% mapivt@0.50:0.95, improving upon yolov12 from 34.3% to 36.1% and from 31.8% to 33.1%, respectively. - **Candidate:** our proposed tdnet achieves the best results across all

major components. it reaches 36.1% mapivt@0.5 and 33.1% mapivt@0.50:0.95, improving upon yolov12 from 34.3% to 36.1% and from 31.8% to 33.1%, respectively.

8. Parameter-efficient framework for surgical action triplet recognition

URL: [View paper](#)

Brief Assessment

Parameter-efficient Triplet[15] focuses on parameter-efficient model architectures for surgical triplet recognition on CholecT50, not on developing evaluation toolkits or benchmarking methodologies for surgical triplet detection.

9. Surgical activity triplet recognition via triplet disentanglement

URL: [View paper](#)

Brief Assessment

Activity Triplet Disentanglement[19] focuses on a triplet disentanglement framework for surgical action triplet recognition, not on developing evaluation toolkits or benchmarking infrastructure for surgical triplet detection.

10. Rendezvous in time: an attention-based temporal fusion approach for surgical triplet recognition

URL: [View paper](#)

Brief Assessment

Rendezvous Time[18] focuses on temporal modeling for surgical triplet recognition rather than evaluation toolkits or benchmarking infrastructure. The paper does not discuss evaluation toolkit development or benchmark creation.

Appendix: Text Similarity Detection

Textual similarity detection checked 25 papers and found 4 similarity segment(s) across 2 paper(s).

The following **2 paper(s)** were detected to have high textual similarity with the original paper. These may represent different versions of the same work, duplicate submissions, or papers with substantial textual overlap. Readers are advised to verify these relationships independently.

1. ProstaTD: A Large-scale Multi-source Dataset for Structured Surgical Triplet Detection

Detected in: Core Task (sibling), Contribution: [contribution_1](#), Contribution: [contribution_2](#), Contribution: [contribution_3](#)

△ **Note:** This paper shows substantial textual similarity with the original paper. It may be a different version, a duplicate submission, or contain significant overlapping content. Please review carefully to determine the nature of the relationship.

2. CholecTriplet2021: A benchmark challenge for surgical action triplet recognition

Detected in: Contribution: [contribution_3](#)

△ **Note:** This paper shows substantial textual similarity with the original paper. It may be a different version, a duplicate submission, or contain significant overlapping content. Please review carefully to determine the nature of the relationship.

References

- [0] ProstaTD: Bridging Surgical Triplet from Classification to Fully Supervised Detection [View paper](#)
- [1] Guardian: Guarding against uncertainty and adversarial risks in robot-assisted surgeries [View paper](#)
- [2] Artificial intelligence and automation in endoscopy and surgery [View paper](#)
- [3] ProstaTD: A Large-scale Multi-source Dataset for Structured Surgical Triplet Detection [View paper](#)
- [4] Pixel-wise recognition for holistic surgical scene understanding [View paper](#)
- [5] Triple-console robotic telesurgery: first impressions and future impact [View paper](#)
- [6] A deep learning framework for surgery action detection [View paper](#)
- [7] Surgical action triplet recognition via triplet disentanglement [View paper](#)
- [8] V05-12 SINGLE PORT ROBOT-ASSISTED RADICAL PROSTATECTOMY WITH PERITONEAL FLAP INTERPOSITION IN SETTING OF PRIOR TOTAL â€¦ [View paper](#)
- [9] SAR-RARP50: Segmentation of surgical instrumentation and Action Recognition on Robot-Assisted Radical Prostatectomy Challenge [View paper](#)
- [10] Estimating surgical urethral length on intraoperative robot-assisted prostatectomy images using artificial intelligence anatomy recognition [View paper](#)
- [11] Towards holistic surgical scene understanding [View paper](#)
- [12] A Dataset and Benchmark for Robot-Assisted Radical Prostatectomy With Lymphadenectomy in Surgical Workflow Understanding [View paper](#)
- [13] A Dataset for Robot-assisted Radical Prostatectomy with Lymphadenectomy in Surgical Workflow Understanding [View paper](#)
- [14] TriQuery: A Query-Based Model for Surgical Triplet Recognition. [View paper](#)
- [15] Parameter-efficient framework for surgical action triplet recognition [View paper](#)
- [16] Temset-24k: Densely annotated dataset for indexing multipart endoscopic videos using surgical timeline segmentation [View paper](#)
- [17] Instrument-tissue-guided surgical action triplet detection via textual-temporal trail exploration [View paper](#)
- [18] Rendezvous in time: an attention-based temporal fusion approach for surgical triplet recognition [View paper](#)
- [19] Surgical activity triplet recognition via triplet disentanglement [View paper](#)
- [20] CurConMix: A Curriculum Contrastive Learning Framework for Enhancing Surgical Action Triplet Recognition [View paper](#)
- [21] CholecTriplet2021: A benchmark challenge for surgical action triplet recognition [View paper](#)
- [22] Why deep surgical models fail?: Revisiting surgical action triplet recognition through the lens of robustness [View paper](#)
- [23] MT4MTL-KD: A Multi-Teacher Knowledge Distillation Framework for Triplet Recognition [View paper](#)
- [24] Surgical video workflow analysis via visual-language learning [View paper](#)
- [25] â€œDeep-Ontoâ€ network for surgical workflow and context recognition [View paper](#)
- [26] Frame Selection Methods to Streamline Surgical Video Annotation for Tool Detection Tasks [View paper](#)
- [27] Grounding Surgical Action Triplets with Instrument Instance Segmentation: A Dataset and Target-Aware Fusion Approach [View paper](#)
- [28] Surgical Triplet Recognition via Diffusion Model [View paper](#)
- [29] Web based Object Annotation Tool using a Triplet-ReID Sorting Approach [View paper](#)