Sub-Resolution Monocyte Detection in Time-Lapse MRI: A Benchmark

Robin Rexeisen[1][0009−0001−0744−5315], Asli Havlas[2][0009−0000−6402−7112], Enrica Wilken[2][0000−0003−4115−3831], Lydia Wachsmuth[2][0000−0003−1684−7196], Anne Helfen[2][0000−0001−5894−2365], Max Masthoff[2][0000−0001−5338−9193], Cornelius Faber[2][0000−0001−7683−7710], and Xiaoyi Jiang[1][0000−0001−7678−9528]

[1] Department of Computer Science, University of Münster, Münster, Germany
{r.rexeisen,xjiang}@uni-muenster.de
[2] Clinic of Radiology, University of Münster, Münster, Germany
a.havlas@gmx.de, Enrica.Wilken@unimedizin-mainz.de,
{wachslyd,anne.helfen,max.masthoff,faberc}@uni-muenster.de

Abstract. Detecting individual immune cells in time-lapse magnetic resonance imaging (MRI) enables the non-invasive study of inflammatory processes. However, automated detection of contrast-enhanced monocytes remains highly challenging due to their extremely small spatial footprint (2–6 pixels), heterogeneous appearance, temporal variability, and ambiguity with confounding phenomena such as micro-bleeding. In this work, we introduce the first publicly available time-lapse MRI dataset for contrastive monocyte detection. The dataset contains carefully annotated sequences capturing substantial appearance variability and realistic clinical artifacts. We establish baseline performance using state-of-the-art object detection architectures and demonstrate that standard detectors perform suboptimally in the sub-resolution regime, especially under domain shift. Furthermore, we systematically analyze the contribution of key components of the YOLO architecture by progressively pruning its detection heads, feature pyramid levels, and loss terms. Our results reveal which architectural elements are essential for reliable detection of ultra-small objects, with a focus on domain generalization. We believe this dataset and analysis provide a benchmark for future research in microscopic structure detection in medical imaging.

Keywords: Time-Lapse MRI · Immune Cell Detection · Sub-Resolution Monocyte Detection · Benchmark.

1 Introduction

The study of immune cells is crucial for inflammation and cancer research [8,9,15]. However, techniques such as intravital microscopy, though widely used, are invasive and may activate the host-defense system, compromising organism integrity. Their limited field of view also restricts systemic cell migration studies [14]. In contrast, time-lapse MRI with iron oxide nanoparticle (ION) labelling enables

non-invasive investigation of single circulating monocytes, revealing behavioral changes in response to inflammatory stimuli [16].

Current studies rely on the manual detection of cells [1], which appear as small, ambiguous, hypointense spots. This tedious, error-prone, and time-intensive task dramatically slows down research. Typical cell detection algorithms, developed primarily for microscopic images [4], are ill-suited for MRI due to differences in resolution, quality, and the sub-resolution appearance of these spots. This challenge defines a new problem class in MRI: *sub-resolution detection (SRD) of immune cells.*

At present, there is no publicly available dataset that permits reproducible benchmarking of monocyte detection methods in time-lapse MRI. Moreover, the absence of standardized protocols hampers systematic progress, making it difficult to compare results across studies. In this work, we introduce the first publicly available 4D time-lapse MRI dataset for sub-resolution monocyte detection. Beyond releasing data, we establish the first standardized benchmark and analyze detection design choices specific to this extreme small-object regime. **Our contributions** are: 1) A public 4D MRI dataset for sub-resolution monocyte detection with 124 recordings and 115k+ annotations (Section 3). 2) A standardized benchmark protocol with fixed volume-level split, an evaluation metric tailored to 2–6 pixel objects, and reproducible deep learning baselines (Section 4). 3) An empirical analysis of architectural and loss-design choices near the practical limits of MRI spatial resolution (Section 5). We release this dataset[3] and source code[4] for all experiments alongside this paper.

2 Related Work

Small object detection. Small object detection typically targets objects $\leq 32 \times 32$ pixels, where contour and texture cues usually remain informative [19]. In contrast, monocytes in our data span only 2–6 pixels, approaching MRI's spatial resolution limit, where objects appear as subtle intensity perturbations rather than shape-defined entities. Infrared Small Target Detection and Segmentation (IRSTDS), used in applications like missile interception for identifying a small number of targets, shares similar challenges with our dataset, such as low-contrast, sub-resolution targets on heterogeneous backgrounds [29]. Modern networks such as YOLOv8n-p2 variants have shown success here [28]. Our dataset, however, distinguishes by inhabiting more targets per slice within more complex MR scenes. This places sub-resolution monocyte detection in a unique and more demanding regime.

Cell detection. Cell detection, commonly applied in time-lapse microscopy, typically handles larger recorded cells than in our task. Common challenges such as cell contact and overlap [6,17,18] are absent in our application. The Particle Tracking subtask can target 100–1000 point-sized objects such as single

[3] `https://www.uni-muenster.de/PRIA/en/forschung/index.shtml`

[4] `https://zivgitlab.uni-muenster.de/ag-pria/sub-resolution-experiments`

molecules [7]. Despite the comparable target sizes, these recordings feature more homogeneous backgrounds and a higher temporal resolution.

Sub-resolution cell detection in MRI. To our knowledge, Mori et al. [20] first proposed an automated method for detecting sub-resolution immune cells in MRI data, using four-neighbor connected-component labeling on thresholded brain regions segmented via the watershed algorithm. Tashita et al. [25] employed the Hungarian algorithm and an SVM to classify voxel candidates, enhancing accuracy with temporally averaged intensity volumes. Afridi et al. [2] detected hypointense spots and homogeneous neighbors using fine and coarse superpixels, exploiting their intensity differences with a partition-based Bayesian classifier. Later, they compared MLP and CNN for classifying superpixel-based extracted patches of candidate regions, finding CNNs superior [3]. However, all these studies rely on classical methods and private datasets, hindering comparability. Also, the sole deep learning approach is not tailored to detection.

3 Dataset for Sub-Resolution Monocyte Detection

Acquisition. The in vivo time-lapse MRI of mice brain originated from female BALB/c mice. The animal experiments were carried out according to local animal welfare guidelines and were approved by local authorities (ID: 81–02.04.2020.A194). Cells were labeled in vivo by i.v. injection of 3mL per kg body weight Ferucarbotran (Resovist, Bayer AG, Leverkusen, Germany; 0.5mmol Fe/mL) via the tail vein. The recordings were performed 24h after labeling. Time-lapse MRI of the brain was performed on a 9.4 T Biospec (Bruker Biospin, Ettlingen, Germany) following the protocol described in [16]. All annotations were performed by a single experienced biologist following the labeling protocol of our previous work [5,26,27].

Dataset statistics and characteristics. The resulting dataset comprises 124 4D recordings in tif format, each containing 20 frames and 38 axial slices with a side length between 142–180 pixels. Each recording originates from a distinct mouse. In total, there are 94,240 axial slices, with monocytes found in 49,821 (52.78%) of them. We generated 115,770 annotations across all recordings, encoding the x and y coordinates of detected monocytes. A single axial slice can accommodate up to 22 monocytes, corresponding to a maximum of 256 per slice sequence. This yields up to 363 monocytes within a volume at a single time-point, and potentially 5,409 across an entire 4D MRI recording. Figure 1 presents representative examples with labeled monocytes visible as hypointense spots.

The number of annotations per slice follows an exponential distribution (Figure 2a). The spatial distance between annotations within a slice approximates a Rayleigh distribution (Figure 2c). The distances between annotations in consecutive frames exhibit a mixed distribution (Figure 2d). This distribution combines a Rayleigh-like pattern, as shown in Figure 2c, with a second distribution that we approximate as a folded normal. Considering the Rayleigh distribution to rep-

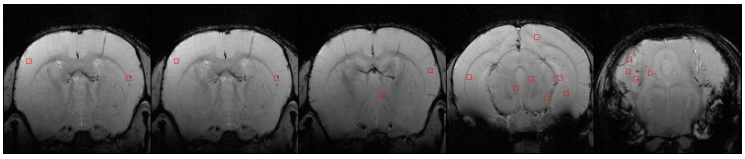

Fig. 1: Randomly sampled slices from a volume within our dataset, showcasing time-lapse MRI of a murine brain with ION-labeled monocytes. Annotation is visualized as red fixed-size bounding boxes around the annotated locations.

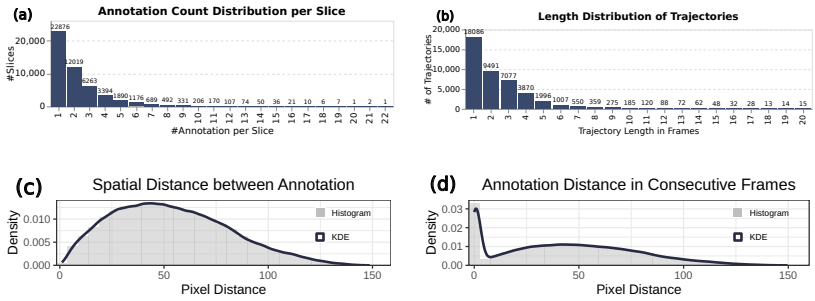

Fig. 2: Statistical properties of the presented dataset.

resent the general distance between independent annotations, the narrow folded normal distribution may reflect the movement behavior of cells.

This analysis enables us to define a principled distance threshold to narrow down linking candidates for trajectory construction. Fitting this distribution assumption to the data indicates that 99% of cell movements are within a 3-pixel distance, following a folded normal distribution. The probability of observing separate cells within this radius in subsequent frames is approximately 0.2%. The distribution of trajectory lengths, determined by this threshold, follows another exponential distribution (see Figure 2b). To resolve ambiguity when an annotation had multiple possible successors, we selected the candidate yielding the longest trajectory.

Challenges for the detection task. Detection in this extreme regime fundamentally differs from conventional small-object detection due to pixel-level discretization, which renders spatial patterns and structural features ineffective. Low signal-to-background contrast is further confounded by similar hypointensities (e.g., blood, air, or anatomical structures). Prolonged recording times per frame and monocyte movement introduce motion blur and low-contrast signal, ultimately reducing the signal-to-noise ratio. As illustrated in Figure 2a, multiple monocytes are possibly present within a single image, up to 22 in our dataset. These cells appear sporadically and frequently disappear within a few frames, which limits temporal trajectory analysis, particularly for prediction refinement. Figure 3 summarizes these challenges.

(a) (b) (c)

Fig. 3: Key challenges in SRD. (a) High-magnification view of a monocyte and background, highlighting the difficulty in resolving hypointense spots due to their small size. (b) Highly magnified monocytes illustrating the low contrast issue. (c) Slice sequence excerpt with annotated red boxes, showing the fleeting temporal presence of targets.

Finally, we acknowledge that manual annotation may result in occasional mislabeling or omissions. Nevertheless, we believe our dataset provides a valuable benchmark for future research on monocyte detection in time-lapse MRI.

4 Standardized Benchmark

Data split and k-Fold cross-validation protocol. To ensure reproducibility and comparability of future methods, we define a standardized benchmark for sub-resolution immune cell detection. Dataset splitting is performed at the volume level to build a test set (25 volumes from distinct mice with 18,000 slices) separate from a k-Fold set (99 volumes and 56,949 slices) in order to examine the domain shift capability.

Experimental setup. To establish the first baseline performance on this dataset, we evaluated Faster-RCNN [22] with a ResNet50 backbone, as well as YOLOv8, YOLO11, and YOLO26 [10,11,12,21,24]. To ensure a fair comparison, we trained all networks from scratch and disabled data augmentation. We set $k = 5$ for all experiments in this section to assess the empirical robustness of the training process. We processed the data slice-wise, with bounding boxes placed centered at the manual annotation, to enable the usage of these networks. These boxes are set to a 10-pixel edge size, as the network failed to converge at far smaller sizes. All networks were trained until no further improvements in the loss function were observed. YOLO's training parameters used default values from the ultralytics framework [13], employing a learning rate of 0.01 with MuSGD. Faster-RCNN was trained using SGD with a learning rate of 0.01.

Spatio-temporal context aggregation. In our experiments, we investigated the effect of additional input channels by stacking related slices. First, we incorporated two adjacent axial slices. Our reasoning was that while primary brain structures are visible in neighboring slices, the slice thickness generally prevents a monocyte to inhabit all three simultaneously, which contributes to their distinguishability. Alternatively, we incorporated the temporal dimension by including the temporally average of the complete slice sequence, inspired by [25]. Due to their short appearance, monocytes are averaged out like noise, leaving persistent

#	Network	R	BF	N	Avg	AUC (k-Fold)	AUC (Testset)
1	YOLO11n-p2	✔	✗	✔	✔	0.8725 ± 0.003	0.7049 ± 0.012
2	YOLOv8n-p2	✔	✗	✔	✔	0.8685 ± 0.007	0.6978 ± 0.026
3	YOLOv8n-p2	✔	✗	✗	✔	0.8667 ± 0.012	0.6818 ± 0.019
4	YOLOv8n-p2	✔	✔	✔	✔	0.8665 ± 0.002	0.7109 ± 0.024
5	YOLOv8n-p2	✗	✗	✔	✔	0.8630 ± 0.003	0.6523 ± 0.017
6	YOLO26n-p2	✔	✗	✔	✔	0.8607 ± 0.010	0.7213 ± 0.019
7	YOLOv8n-p2	✗	✗	✗	✔	0.8590 ± 0.006	0.6496 ± 0.024
8	YOLOv8n-p2	✗	✗	✗	✗	0.8213 ± 0.020	0.2803 ± 0.010
9	YOLOv8n-p2	✔	✗	✔	✗	0.8190 ± 0.009	0.2776 ± 0.007
10	YOLOv8n-p2	✗	✗	✔	✗	0.7992 ± 0.061	0.2863 ± 0.021
11	YOLOv8n	✗	✗	✔	✔	0.7808 ± 0.056	0.6563 ± 0.034
12	Faster-RCNN	✗	✗	✔	✔	0.6514 ± 0.028	0.5333 ± 0.018

Table 1: Baseline results for the Monocyte Detection Task. "R" denotes Rigid Registration, "BF" Bias Field Correction, "N" inclusion of neighboring slices as input, and "Avg" inclusion of the temporal average.

background data and thus supports the network's ability to distinguish targets from other hypointense spots. We further explored combinations of these variations. Additionally, we examined the effect of rigid image registration and bias field correction on the performance, which compensates for subject movement and reduces intensity variations stemming from the MR hardware, respectively.

Performance metric. Conventional IoU-based detection metrics become unstable in the sub-resolution regime (dominated by discretization effects rather than semantic localization quality). For objects spanning 2–6 pixels, a 1-pixel localization deviation can result in large IoU variations, making IoU unnecessarily opaque in its interpretability. We therefore define detection correctness in terms of distance tolerance. Specifically, we employ a Maximum Bipartite Matching Algorithm to associate each annotation with at most one prediction within a specific distance. This results in False Positives (FP), False Negatives (FN), and True Positives (TP), from which Recall and Precision is derived for different probability thresholds for the prediction, which ultimately leads to the Area Under the Curve (AUC), which is our final performance. We select a 4-pixel tolerance in our experiments, as we empirically found AUC improvements beyond this radius to be negligible across multiple trained networks.

Baseline performance. According to Table 1, Faster-RCNN demonstrates significantly lower performance on this dataset compared to its YOLO counterpart, in which the P2 variants are superior in the k-Fold evaluation. Image registration and bias field correction show minimal impact on the k-Fold cross-validation performance, but improve generalization on the test set, suggesting that these preprocessing steps help mitigate domain shift. Table 1 confirms that domain shift is a significant problem, as reflected in the test set values, a common challenge in medical machine learning [23]. Notably, incorporating the temporal average yields a substantial performance boost there, apart from improvements

#	CIoU	DFL	IoU	Dist	AR	Variant	BD	Pool	NU	ND	AUC (k-Fold)	AUC (Testset)
1	7.5	1.5	✔	✔	✔	**P1**	5	3	**4**	**4**	0.8722 ± 0.002	0.6984 ± 0.002
2	7.5	1.5	✔	✔	✔	**P1**	5	3	**4**	**1**	0.8703 ± 0.004	0.6878 ± 0.012
3	7.5	1.5	✔	✔	✔	**P1**	5	3	**4**	**0**	0.8638 ± 0.007	0.7086 ± 0.032
4	7.5	1.5	✔	✔	✗	P2	5	3	3	3	0.8579 ± 0.001	0.7120 ± 0.002
5	7.5	1.5	✔	✔	✔	P2	5	**2**	3	3	0.8571 ± 0.003	0.7221 ± 0.010
6	7.5	1.5	✔	✗	✔	P2	5	3	3	3	0.8569 ± 0.007	0.7107 ± 0.023
7	7.5	1.5	✔	✔	✔	P2	5	3	3	**1**	0.8542 ± 0.006	0.6949 ± 0.028
8	7.5	1.5	✔	✔	✔	P2	5	3	3	3	0.8477 ± 0.006	0.7086 ± 0.035
9	7.5	1.5	✔	✔	✔	P2	**3**	3	**1**	**1**	0.8465 ± 0.003	0.7167 ± 0.008
10	**0.0**	1.5	✗	✗	✗	P2	5	3	3	3	0.8460 ± 0.002	0.7323 ± 0.002
11	7.5	1.5	✔	✔	✔	P2	5	**1**	3	3	0.8444 ± 0.021	0.7264 ± 0.015
12	7.5	1.5	✔	✔	✔	P2	5	**2**	3	**1**	0.8417 ± 0.011	0.7385 ± 0.013
13	7.5	1.5	✗	✔	✔	P2	5	3	3	3	0.8332 ± 0.019	0.7424 ± 0.013
14	7.5	**0.0**	✔	✔	✔	P2	5	3	3	3	0.8295 ± 0.050	0.7127 ± 0.015

Table 2: Performance of YOLOv8n modifications (bold) against the default P2 setup (gray row). The first five columns define the loss terms: CIoU (consisting of IoU, normalized center distance "Dist," and aspect ratio term "AR") and DFL. Next architectural changes: "BD" (backbone down-steps), "Pool" (SPPF max-pooling operations), "NU" (neck up-sampling), and "ND" (neck down-sampling).

on the evaluation set. Including neighboring slices, in contrast, had a minimal impact on generalization performance. This demonstrates the crucial role of a smoothed, averaged reference image for generalization across the domain, suggesting that the network otherwise overfits to the specific noise patterns present in the training data. YOLO11n-p2 and YOLO26n-p2 show comparable k-Fold performance to YOLOv8n-p2, with YOLO26n-p2 demonstrating improved generalization capability.

5 Key Architectural Components for SRD

Following the initial comparative study, we observed that YOLOv8n-p2 is a representative sub-resolution detector in both k-Fold validation and the test set, demonstrating competitive performance under the tested variants within the domain and under domain-shift conditions. Additionally, its effectiveness has already been established in other minuscule object detection tasks [28].

We therefore use this network to analyze architectural and loss-design choices, systematically investigating essential and redundant components for objects spanning few pixels. By modifying the backbone, neck, and localization losses, we isolate performance-governing factors near MRI spatial resolution limits. We also explore a rarely implemented P1 variant with an additional concatenated upsampling layer. Our experiments incorporate all additional inputs from the previous section and apply rigidly registered and bias field-corrected preprocessing, as this approach improves generalization under domain shift while simultaneously

mitigating side effects irrelevant to the sub-resolution detection task. Since this study specifically aims to target architectural components for domain-shift robustness, we use 3-Fold cross-validation; the smaller training proportion per fold better highlights which components contribute to generalization. We present the results in table 2.

Architectural study. The P1 variant outperforms P2 on the k-Fold set, yet shows no significant advantage on the test set, suggesting that the extra fine-scale head mainly focus on domain-specific patterns. Reducing the number of detection heads for coarser resolutions yields no clear effect, suggesting these coarser heads are redundant. Decreasing the number of backbone layers yields largely similar performance on both sets, suggesting that deeper layers provide no meaningful benefit for sub-resolution detection. Reducing the number of pooling operations in the SPPF layer has little effect on k-Fold performance but consistently improves test set generalization, suggesting that fewer pooling operations may help mitigate domain shift.

Loss study. The role of Complete Intersection over Union (CIoU) is nuanced. While CIoU can compensate for the absence of Distribution Focal Loss (DFL) by achieving comparable generalization with a small decline in the k-Fold performance, replacing CIoU entirely with DFL barely affects k-Fold accuracy, but improves test set generalization and reduces the standard deviation notably. This suggests that the components of CIoU, particularly the aspect ratio term, which provides no discriminative signal for our fixed-sized targets, introduce interference that limits cross-domain generalization. The distance and Intersection over Union (IoU) terms individually, however, demonstrate minor effects.

Key components. In conclusion, the loss formulation and pooling configuration are the key components for effective sub-resolution detection. Replacing CIoU with DFL and reducing pooling operations in SPPF improve cross-domain generalization, while architectural choices such as backbone depth and detection heads show limited impact.

Qualitative evaluation. We evaluated typical error cases for the first network in Table 2 on the evaluation set (Figure 4). The network repeatedly struggled with blurry monocyte signals or with low signal-to-background situations. It also confused temporal fluctuations with monocytes. Occasionally, cells couldn't be separated from tissue structures. Despite annotations being restricted to the brain region, the network can fail to restrict results to this area. Finally, ringing artifacts are also found to correlate with increased network failure.

6 Conclusion

This work establishes the first standardized benchmark for sub-resolution immune cell detection in MRI. By releasing a large-scale annotated dataset with an evaluation protocol and reproducible baselines, we aim to advance automated single-cell MRI analysis. Our experiments show that state-of-the-art detection layers can achieve strong baseline performance, yet domain shift remains a ma-

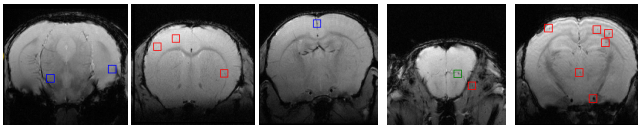

Fig. 4: Failure of the trained YOLO network (red=FP, blue=FN, green=TP). From left to right: signal blur, temporal fluctuations, monocytes near tissue structures, out-of-brain region detections, and confusion due to ringing artifacts.

jor limitation. Incorporating temporally averaged information proves crucial for this role, while DFL and reduced SPPF pooling are additional key drivers of cross-domain performance. In contrast, CIoU loss hinders generalization.

Overall, framing sub-resolution immune cell detection as a benchmark supports progress at MRI's spatial limits and paves the way for temporally aware models, physics-informed approaches, and domain-adaptive strategies for clinical applications in future.

Acknowledgments. This work was supported by the DFG (German Research Foundation) via CRC 1450 "InSight" (project number 431460824), including projects B01, B02, and Z01, the CiM-IMPRS Graduate School, and project 549903517. Calculations were performed on the HPC cluster PALMA II (University of Münster), funded by the DFG (INST 211/667-1).

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Book of Abstracts ESMRMB 2021 Online 38th Annual Scientific Meeting. Magnetic Resonance Materials in Physics, Biology and Medicine **34**(S1), 1–204 (2021)
2. Afridi, M.J., Liu, X., Shapiro, E., Ross, A.: Automatic in vivo cell detection in MRI. In: MICCAI, pp. 391–399 (2015)
3. Afridi, M.J., Ross, A., Liu, X., Bennewitz, M.F., Shuboni, D.D., Shapiro, E.M.: Intelligent and automatic in vivo detection and quantification of transplanted cells in MRI. Magnetic Resonance in Medicine **78**(5), 1991–2002 (2017)
4. Ahrens, E.T., Bulte, J.W.M.: Tracking immune cells in vivo using magnetic resonance imaging. Nature Reviews Immunology **13**(10), 755–763 (2013)
5. Armstrong, M., Wilken, E., Freppon, F., Masthoff, M., Faber, C., Xiao, D.: Dynamic cell tracking using time-lapse MRI with variable temporal resolution Cartesian sampling. Magnetic Resonance in Medicine **90**(6), 2443–2453 (2023)
6. Chen, M.: Cell tracking in time-lapse microscopy image sequences. In: Computer Vision for Microscopy Image Analysis, pp. 101–129 (2021)
7. Chenouard, N., et al.: Objective comparison of particle tracking methods. Nature Methods **11**(3), 281–289 (2014)
8. Grivennikov, S.I., Greten, F.R., Karin, M.: Immunity, inflammation, and cancer. Cell **140**(6), 883–899 (2010)
9. Helfen, A., et al.: Target-specific imaging of cathepsin and S100A8/A9 reflects specific features of malignancy and enables estimation of tumor malignancy. Molecular Imaging and Biology **22**(1), 66–72 (2020)
10. Jocher, G., Chaurasia, A., Qiu, J.: Ultralytics yolov8 (2023), `https://github.com/ultralytics/ultralytics`
11. Jocher, G., Qiu, J.: Ultralytics yolo11 (2024), `https://github.com/ultralytics/ultralytics`
12. Jocher, G., Qiu, J.: Ultralytics yolo26 (2026), `https://github.com/ultralytics/ultralytics`
13. Jocher, G., Qiu, J., Chaurasia, A.: Ultralytics YOLO (Jan 2023), `https://github.com/ultralytics/ultralytics`
14. Kircher, M.F., Gambhir, S.S., Grimm, J.: Noninvasive cell-tracking methods. Nature Reviews Clinical Oncology **8**(11), 677–688 (2011)
15. Mantovani, A., Allavena, P., Sica, A., Balkwill, F.: Cancer-related inflammation. Nature **454**(7203), 436–444 (2008)
16. Masthoff, M., et al.: Resolving immune cells with patrolling behaviour by magnetic resonance time-lapse single cell tracking. EBioMedicine **73**, 103670 (2021)
17. Maška, M., et al.: A benchmark for comparison of cell tracking algorithms. Bioinformatics **30**(11), 1609–1617 (2014)
18. Maška, M., et al.: The Cell Tracking Challenge: 10 years of objective benchmarking. Nature Methods **20**(7), 1010–1020 (2023)
19. Mirzaei, B., Nezamabadi-pour, H., Raoof, A., Derakhshani, R.: Small object detection and tracking: A comprehensive review. Sensors **23**(15), 6887 (2023)
20. Mori, Y., Chen, T., Fujisawa, T., Kobashi, S., Ohno, K., Yoshida, S., Tago, Y., Komai, Y., Hata, Y., Yoshioka, Y.: From cartoon to real time MRI: In vivo monitoring of phagocyte migration in mouse brain. Scientific Reports **4**(1), 6997 (2014)
21. Ramos, L.T., Sappa, A.D.: A decade of you only look once (YOLO) for object detection: A review. IEEE Access **13**, 192747–192794 (2025)
22. Ren, S., He, K., Girshick, R., Sun, J.: Faster R-CNN: Towards real-time object detection with region proposal networks. In: NIPS (2015)

23. Richiardi, J., Ravano, V., Molchanova, N., Gordaliza, P.M., Kober, T., Cuadra, M.B.: Domain shift, domain adaptation, and generalization. In: Trustworthy AI in Medical Imaging, pp. 127–151. Elsevier (2025)
24. Sapkota, R., et al.: YOLO advances to its genesis: a decadal and comprehensive review of the you only look once (YOLO) series. Artificial Intelligence Review **58**(9), 274 (2025)
25. Tashita, A., Kobashi, S., Nii, M., Mori, Y., Yoshioka, Y., Hata, Y.: Macrophage learning-tracking algorithm in time-lapse MR images. In: International Conference on Machine Learning and Cybernetics (ICMLC). pp. 421–426 (2016)
26. Wilken, E., Havlas, A., Masthoff, M., Moussavi, A., Boretius, S., Faber, C.: Radial compressed sensing imaging improves the velocity detection limit of single cell tracking time-lapse MRI. Magnetic Resonance in Medicine **91**(4), 1449–1463 (Apr 2024)
27. Wilken, E., Havlas, A., Wachsmuth, L., Masthoff, M., Diwoky, C., Faber, C.: Boosting the velocity detection limit of 3D single-cell tracking time-lapse MRI by balanced SSFP imaging. Magnetic Resonance in Medicine **94**(3), 1152–1165 (2025)
28. Yang, J., Liu, S., Wu, J., Su, X., Hai, N., Huang, X.: Pinwheel-shaped convolution and scale-based dynamic loss for infrared small target detection. In: AAAI. pp. 9202–9210 (2025)
29. Zhao, M., Li, W., Li, L., Hu, J., Ma, P., Tao, R.: Single-frame infrared small-target detection: A survey. IEEE Geoscience and Remote Sensing Magazine **10**(2), 87–119 (2022)