RELAXED DISTRIBUTION-WISE CONSISTENCY: LEARNING AFFINITY ORDER FOR SEMI-SUPERVISED MEDICAL IMAGE SEGMENTATION

Anonymous authors
Paper under double-blind review

ABSTRACT

Semi-supervised medical image segmentation learns from pseudo-labels produced by an imperfect teacher, and its supervision faces a tension between reliability and richness. Pseudo-label filtering pursues reliability, but discards the boundary pixels where supervision is most needed; inter-pixel supervision pursues richness and reaches these pixels, but aligns the teacher's noisy values, such as affinity distributions, in a value-alignment manner. Since noisy values may still preserve the correct order, we propose Relaxed Distribution-wise Consistency (RDC), which relaxes inter-pixel alignment from values to order and combines it with filtered pixel-level pseudo-labels. RDC is a simple and lightweight addition to the standard teacher–student framework: each pixel is represented by affinities to well-separated reference pixels (agents), and the affinity order is aligned between teacher and student through a differentiable Kendall's τ. Our analysis shows that, under teacher noise, value alignment drifts and can even reverse the order of two agents, whereas order alignment keeps the direction of the expected update of each pairwise affinity order and only weakens it; we further bound the fraction of inverted affinity orders. On ACDC, LA, Pancreas-NIH, and Synapse, RDC achieves state-of-the-art performance with the largest gains on the most difficult organs, and consistently improves GapMatch, GraphCL, VQ-Seg, and the SAM-based ESDE as a plug-in. With the agents fixed, every order alignment outperforms every value alignment.

1 INTRODUCTION

Medical image segmentation, which aims to delineate internal structures such as organs and lesions in computed tomography (CT) or magnetic resonance imaging (MRI), has achieved remarkable progress with recent advances in deep neural networks (Cao et al., 2022), enabling applications such as autonomous surgical robotics (Haidegger, 2019) and augmented reality-assisted surgery (Birlo et al., 2022). Compared with natural images, medical images often show weak contrast between neighboring tissues, and many target structures are small or thin, so the most difficult decisions concentrate near boundaries (Cao et al., 2022). Making these decisions reliably requires dense pixel-wise annotations, which are labor-intensive to collect and demand clinical expertise, motivating considerable work on semi-supervised medical image segmentation (SSMIS) (Yu et al., 2019). Since only limited labeled medical images are accessible, how to develop reliable supervision signals that fully exploit substantial unlabeled data to improve the model's generalization performance remains extremely challenging.

To this end, the consistency regularization paradigm (Chen et al., 2021; Liu et al., 2022; Chi et al., 2024; Wu et al., 2021) has become dominant in this field: a teacher network, updated as an exponential moving average of the student, generates pseudo-labels on weakly augmented images, and the student is trained to align with them *at the individual pixel level* under strong augmentation (Figure 1a). However, the teacher is itself trained with limited annotations, and its pseudo-labels are least reliable exactly where segmentation is most difficult, near boundaries and on small structures such as the myocardium (Cao et al., 2022). Trained against such targets, the student gradually absorbs the teacher's errors in a self-reinforcing manner, an issue known as confirmation bias (Guo et al., 2017).

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

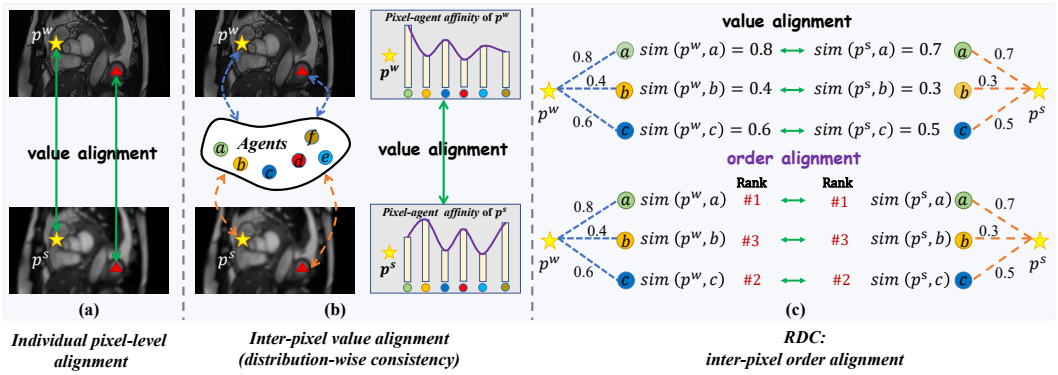


Figure 1: Supervision strategies under an imperfect teacher. **(a)** Individual pixel-level alignment aligns the teacher and student predictions of the same pixel (p^w and p^s), so noise in the pseudo-labels directly enters the student's training signal. **(b)** Inter-pixel value alignment describes each pixel by affinities to reference pixels (agents) and aligns the resulting affinity distributions in a value-alignment manner (Mai et al., 2024), driving the student toward the noisy values computed by the teacher. **(c)** In this example, the student already preserves the teacher's order ($a \succ c \succ b$) with different values. Value alignment still penalizes the student, whereas the order alignment of RDC is already satisfied.

To mitigate this issue, mainstream methods develop various *supervision strategies* upon this teacher–student framework, such as filtering cleaner pseudo-labels (Yu et al., 2019), exploring richer inter-pixel supervision (Wu et al., 2022; Mai et al., 2024), and designing effective teacher–student ensembling (Chen et al., 2021; Wu et al., 2021). One representative strategy keeps the individual pixel-level alignment but *pursues reliability* through pseudo-label filtering, retaining only confident predictions either through a fixed confidence threshold (Chi et al., 2024) or through uncertainty estimated by Monte Carlo dropout (Yu et al., 2019). Such reliability, however, comes at the cost of richness: the discarded pixels concentrate near boundaries, where supervision is most needed, leaving the student least guided where it errs most.

Another representative strategy instead *pursues richer supervision* to better exploit unlabeled data: beyond individual pixel-level alignment, it exploits inter-pixel relations that describe how pixels relate to one another. Such inter-pixel supervision is typically built either among pixels themselves or between pixels and representative references. Contrastive methods (Wu et al., 2022; Wang et al., 2024; You et al., 2022; Zhao et al., 2022) take the former route, pulling together pixels assigned to the same class and pushing apart those assigned to different classes, while RankMatch (Mai et al., 2024) takes the latter, characterizing each pixel by its affinities to a set of representative reference pixels, and enforcing consistency on the resulting affinity distributions between teacher and student, which we term distribution-wise consistency (Figure 1b). More importantly, inter-pixel relations can remain informative where individual pixel predictions are not: the prediction of a boundary pixel may be ambiguous between two adjacent organs, yet this pixel is still clearly more similar to pixels of these organs than to background pixels. Inter-pixel supervision can therefore be imposed on the pixels that filtering leaves behind.

Richness, however, does not guarantee reliability. Like individual pixel-level alignment, existing inter-pixel methods enforce consistency on pairwise similarities (affinity distributions) derived from the teacher's outputs in a *value*-alignment manner (Figure 1a,b). Since the teacher is imperfect, the values are inevitably noisy, making the supervision signals fragile for the student to align with. For contrastive methods, noise further enters through pair construction, since pixels are assigned to positive or negative pairs by the teacher's own predictions. In fact, noisy values may still preserve the correct order, indicating which of two pixels is more similar to a given pixel (*e.g.*, the boundary pixel remains more similar to organ pixels than to background ones). However, value alignment disregards the preserved order and continues to drive the student toward the noisy values, even after the student has learned the correct order (Figure 1c, top). Worse, since the teacher's values fluctuate with augmentation and training, the target that value alignment pursues in expectation is itself displaced, and it can even reverse the order of two reference pixels that the teacher ranks correctly in most views, pulling the student away from the correct order (Section 4). Overall, filtering gains reliability at the cost of richness, whereas inter-pixel supervision gains richness at the cost of

reliability. This leaves the challenge raised at the outset unresolved: *how can supervision from an imperfect teacher be both reliable and rich, so as to better exploit unlabeled data?*

In this paper, we address this challenge by relaxing inter-pixel alignment from *values* to *order*. Since the order of inter-pixel relations is much more robust to teacher noise than the values, an imperfect teacher can still provide reliable supervision through the order alone. Accordingly, the student learns which pixels a given pixel is more similar to, while the similarity values are free to differ from the teacher's. In this way, order supervision enjoys the synergy of the two strategies above, retaining the richness of inter-pixel supervision without sacrificing reliability, and thereby extending supervision to the uncertain pixels that filtering leaves behind. Since the order carries no class semantics, it is combined with the filtered pixel-level pseudo-labels, which continue to supply class semantics on confident pixels. We instantiate this strategy as **R**elaxed **D**istribution-wise **C**onsistency (RDC), which relaxes distribution-wise consistency from values to order (Figure 1c). RDC is a simple and lightweight addition to the standard teacher–student framework with two components: **(i)** *agent construction*, which represents each pixel by affinities to well-separated reference pixels (agents), realized by orthogonal agent selection; and **(ii)** *order-based alignment supervision*, which aligns the affinity order between teacher and student, realized by a differentiable Kendall's τ (Kendall, 1938).

Beyond this intuition, we analyze what the student learns from an imperfect teacher under the two alignments (Section 4). For value alignment, teacher noise displaces the target: the expected update consists of the update toward the teacher's reference affinities plus a drift term, which can push the student against the correct direction even when the teacher's order never changes. For order alignment, teacher noise only scales down the expected update of each pairwise affinity order, so every affinity order keeps its direction and is merely weakened; we further bound the expected fraction of inverted affinity orders, so that a few near-ties leave the supervision of the remaining ones intact. Extensive experiments on four benchmarks (ACDC (Bernard et al., 2018), LA (Xiong et al., 2021), Pancreas-NIH (Roth et al., 2015), and Synapse (Chang et al., 2023)) demonstrate that RDC achieves state-of-the-art performance, with consistent gains as a plug-in to competitive frameworks, including GapMatch (Huang et al., 2025), GraphCL (Wang et al., 2024), VQ-Seg (Yang et al., 2026), and the SAM-based ESDE (Zhao et al., 2026). More importantly, controlled experiments trace these gains to the alignment itself: with the agents fixed and only the alignment changed, every order alignment outperforms every value alignment, and restricting RDC to confident pixels forfeits 1.30 of its 3.24 Dice gain, confirming that order supervision reaches the uncertain pixels that filtering leaves behind.

In summary, our contributions are threefold. **(1) Strategy.** We identify the tension between reliability and richness in supervision from an imperfect teacher, and propose RDC, which relaxes inter-pixel alignment from values to order and thereby extends reliable supervision to the uncertain pixels that filtering leaves behind. **(2) Analysis.** We characterize what the student learns from an imperfect teacher: value alignment drifts and can even reverse the order of reference pixels, whereas order alignment keeps the direction of each pairwise affinity order and only weakens it. **(3) Evidence.** RDC achieves state-of-the-art results on four medical benchmarks and consistent plug-in gains, and controlled experiments isolate the alignment and verify the analysis on trained models.

2 Related Work

Most SSMIS methods build on teacher–student consistency and differ in their supervision strategies: UA-MT (Yu et al., 2019) filters pseudo-labels by uncertainty, confidence thresholds (Sohn et al., 2020; Zhang et al., 2021; Wang et al., 2022; Chen et al., 2023) decide which pseudo-labels reach the student, and BCP (Bai et al., 2023), ABD (Chi et al., 2024), GapMatch (Huang et al., 2025), and VQ-Seg (Yang et al., 2026) design stronger perturbations of inputs, features, or tokens. Inter-pixel supervision complements these pixel-level signals: contrastive methods (Wang et al., 2021; Alonso et al., 2021; Wu et al., 2022; Wang et al., 2024) build positive and negative pairs, ARCO (You et al., 2023) reduces the variance of the contrastive estimator, and RankMatch (Mai et al., 2024) characterizes each pixel by its affinities to agents and enforces consistency on these affinities between teacher and student. RDC builds on this pixel–agent representation and relaxes the alignment from values to order, so that inter-pixel supervision remains reliable under an imperfect teacher; Table 22 summarizes the differences from RankMatch. Foundation models such as SAM (Kirillov et al., 2023) and its medical adaptations (Ma et al., 2024; Zhu et al., 2024) serve as pseudo-label generators or backbones in semi-supervised frameworks (Zhang et al., 2024; Lu et al., 2024; Zhao et al., 2026),

to which RDC applies as a plug-in. Section F discusses related work in detail, including ranking objectives and noise-tolerant losses.

3 METHOD

RDC keeps the standard teacher–student framework and relaxes inter-pixel alignment from values to order. Section 3.1 recalls this framework, Section 3.2 relates each pixel to a set of agents, Section 3.3 introduces the order alignment of RDC alongside the value alignment it relaxes, and Section 3.4 combines order alignment with the filtered pixel-level pseudo-labels.

3.1 PRELIMINARIES

Let $\mathcal{D}^l = \{(\boldsymbol{x}_i^l, \boldsymbol{y}_i^l)\}_{i=1}^{N^l}$ and $\mathcal{D}^u = \{\boldsymbol{x}_i^u\}_{i=1}^{N^u}$ be the labeled and unlabeled sets, with $N^u \gg N^l$. Following standard practice (Yu et al., 2019; Bai et al., 2023; Chi et al., 2024), a student $f_S(\cdot;\boldsymbol{\theta}_S)$ is trained by gradient descent, and a teacher $f_T(\cdot;\boldsymbol{\theta}_T)$ is its exponential moving average (EMA), $\boldsymbol{\theta}_T \leftarrow m\boldsymbol{\theta}_T + (1-m)\boldsymbol{\theta}_S$. Each unlabeled image is fed to the teacher under a weak augmentation $aug(\cdot)$ and to the student under a strong augmentation $\mathcal{A}ug(\cdot)$. For notational simplicity we consider 2D images with HW pixels; the supervised loss is

$$\mathcal{L}_{sup} = \frac{1}{N^l}\sum_{i=1}^{N^l}\frac{1}{HW}\sum_{j=1}^{HW}\ell_{ce}\big(\boldsymbol{y}_{ij}^l, f_S(\boldsymbol{x}_i^l)_j\big), \tag{1}$$

where ℓ_{ce} is the cross-entropy. The teacher assigns each unlabeled pixel a pseudo-label when its confidence $\mathrm{conf}_{ij} = \max_k f_T(aug(\boldsymbol{x}_i^u))_{j,k}$ exceeds a threshold γ,

$$\hat{\boldsymbol{y}}_{ij}^u = \begin{cases} \arg\max_k f_T(aug(\boldsymbol{x}_i^u))_{j,k}, & \mathrm{conf}_{ij} > \gamma, \\ \texttt{ignore_index}, & \text{otherwise}, \end{cases} \tag{2}$$

and the pseudo-label loss is

$$\mathcal{L}_{reg} = \frac{1}{N^u}\sum_{i=1}^{N^u}\frac{1}{HW}\sum_{j=1}^{HW}\ell_{ce}\big(\hat{\boldsymbol{y}}_{ij}^u,\ f_S(\mathcal{A}ug(\boldsymbol{x}_i^u))_j\big). \tag{3}$$

Pixels marked `ignore_index` contribute nothing to $\mathcal{L}_{reg}$, so pixel-level pseudo-supervision reaches only the pixels that survive filtering.

3.2 PIXEL–AGENT RELATIONS

Relating every pair of pixels would require $HW \times HW$ comparisons per image. Following RankMatch (Mai et al., 2024), we instead relate each pixel to a small set of $N \ll HW$ reference pixels, called *agents*.

Affinities and relations. Let $\boldsymbol{F} \in \mathbb{R}^{C\times HW}$ be the pixel embedding before the final classifier, with $\boldsymbol{f}_j \in \mathbb{R}^C$ the embedding of pixel j. Given the indices $\mathcal{I} = \{idx_1, \ldots, idx_N\}$ of the agents, the agents are the embeddings $\boldsymbol{A} = [\boldsymbol{a}_1, \ldots, \boldsymbol{a}_N] \in \mathbb{R}^{C\times N}$ with $\boldsymbol{a}_n = \boldsymbol{F}[:, idx_n]$. The affinities of pixel j and its affinity distribution are

$$z_{jn} = \boldsymbol{f}_j^\top \boldsymbol{a}_n, \qquad \boldsymbol{c}_j = \mathrm{softmax}(\boldsymbol{z}_j) \in \Delta^{N-1}. \tag{4}$$

Each pair of agents (p, q) defines a *relation* of pixel j, whose direction $\mathrm{sign}(z_{jp} - z_{jq})$ indicates which of the two agents the pixel is more similar to. Superscripts w and s denote teacher and student quantities, *e.g.*, $\boldsymbol{z}_j^w = (\boldsymbol{A}^w)^\top \boldsymbol{f}_j^w$ and $\boldsymbol{z}_j^s = (\boldsymbol{A}^s)^\top \boldsymbol{f}_j^s$; an image index i is added when needed.

Agents. We select agents with the orthogonal agent selection of RankMatch (Mai et al., 2024): starting from a random pixel, it repeatedly adds the pixel whose largest cosine similarity to the agents already selected is the smallest, which avoids near-duplicate agents (Algorithm 1, Section D). The indices $\mathcal{I}$ are computed once per image from the teacher embedding, and the agents $\boldsymbol{A}^w$ and $\boldsymbol{A}^s$ are gathered from the teacher and student embeddings at these indices. The two views share one spatial grid: the geometric augmentation is sampled once to produce the weak view, the strong view applies photometric perturbations and CutMix to the same transformed image, and the same CutMix mask is applied on both sides. An index therefore refers to the same location in both views.

3.3 Value Alignment and Order Alignment

Given the teacher and student affinities of a pixel, an alignment specifies what the student should match.

Value alignment. Distribution-wise consistency (DC) asks the student to reproduce the teacher's affinity distribution,

$$\mathcal{L}_{dc} = \frac{1}{N^u} \sum_{i=1}^{N^u} \frac{1}{HW} \sum_{j=1}^{HW} \mathrm{KL}\big(\boldsymbol{c}_{ij}^{w} \,\|\, \boldsymbol{c}_{ij}^{s}\big). \tag{5}$$

RankMatch also enforces consistency on these affinities in a value-alignment manner, which Section 4 formalizes at the level of permutation probabilities (Lemma 1).

Order alignment. RDC relaxes this target to the order of the affinities. Kendall's rank correlation (Kendall, 1938),

$$\tau(\boldsymbol{z}^w, \boldsymbol{z}^s) = \frac{2}{N(N-1)} \sum_{p<q} \mathrm{sign}(z_p^w - z_q^w)\, \mathrm{sign}(z_p^s - z_q^s), \tag{6}$$

is the fraction of relations on which teacher and student agree minus the fraction on which they disagree, and equals 1 exactly when both order the agents identically. We optimize the differentiable surrogate of DiffKendall (Zheng et al., 2023), which replaces the sign function with $g_\alpha(x) = \tanh(\alpha x/2)$ and has also been used for logit distillation from a trained teacher (Guan et al., 2024):

$$\hat{\tau}_\alpha(\boldsymbol{z}^w, \boldsymbol{z}^s) = \frac{2}{N(N-1)} \sum_{p<q} g_\alpha(z_p^w - z_q^w)\, g_\alpha(z_p^s - z_q^s), \tag{7}$$

where the temperature $\alpha > 0$ controls the sharpness and $\hat{\tau}_\alpha \to \tau$ as $\alpha \to \infty$. The RDC loss is

$$\mathcal{L}_{rdc} = \frac{1}{N^u} \sum_{i=1}^{N^u} \frac{1}{HW} \sum_{j=1}^{HW} \Big(1 - \hat{\tau}_\alpha(\boldsymbol{z}_{ij}^w, \boldsymbol{z}_{ij}^s)\Big), \tag{8}$$

with the teacher affinities detached.

Relational pseudo-labels. For a relation (p, q) of a pixel, let $u = z_p^w - z_q^w$ and $v = z_p^s - z_q^s$ be the teacher and student gaps. The loss of the relation, $\ell(u, v) = 1 - g_\alpha(u)\, g_\alpha(v)$, decomposes exactly as

$$\ell(u, v) = (1 - w) + w\big[1 - y\, g_\alpha(v)\big], \qquad y = \mathrm{sign}(u), \quad w = |g_\alpha(u)|. \tag{9}$$

The first term is constant for the student because the teacher is detached. Each relation thus provides a relational pseudo-label y, the direction preferred by the teacher, with a weight w that grows with the teacher margin $|u|$ and saturates at 1. The student term equals $2\,\mathrm{sigm}(-\alpha y v)$, the sigmoid loss, with $\mathrm{sigm}(\cdot)$ the sigmoid function. Its values for the two directions sum to a constant, so it belongs to the symmetric losses, under which risk minimization is tolerant to label noise (Ghosh et al., 2017). RDC therefore trains the student on pairwise relational pseudo-labels with a margin-weighted symmetric loss. The target of each relation is a direction, and the affinity values of the teacher enter only through the weights.

3.4 Joint Objective

RDC combines order alignment with the filtered pixel-level pseudo-labels:

$$\mathcal{L} = \mathcal{L}_{sup} + \mathcal{L}_{reg} + \lambda\, \mathcal{L}_{rdc}, \tag{10}$$

where λ balances the inter-pixel term. The two unlabeled terms supervise different quantities. $\mathcal{L}_{reg}$ supplies class semantics on the pixels where the teacher is confident, and $\mathcal{L}_{rdc}$ supervises the relations of every pixel, including the uncertain pixels that filtering with the threshold γ leaves behind. These pixels concentrate along organ boundaries (Figure 11), and on them order alignment is the only source of unlabeled supervision. Figure 2 illustrates RDC, and Algorithm 2 in Section D details one training iteration. Agents, affinities, and $\mathcal{L}_{rdc}$ are used only during training, so inference is unchanged.

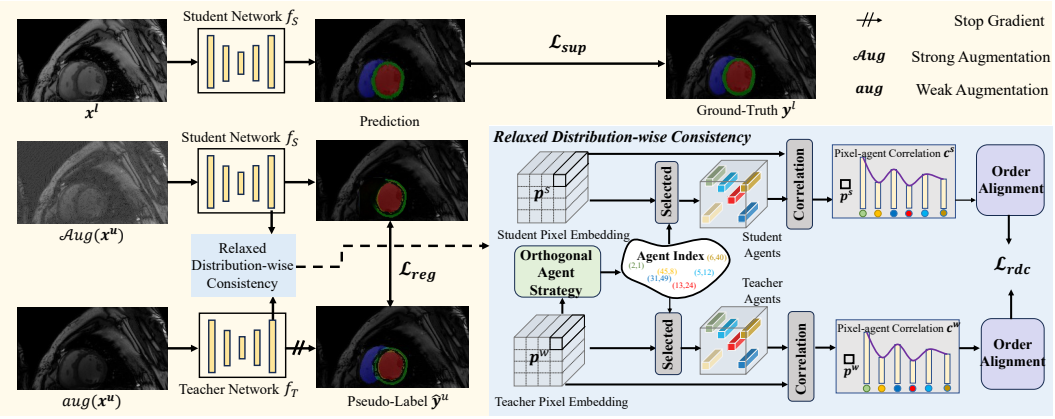


Figure 2: Overall framework of RDC. On top of the standard teacher–student framework with the supervised loss $\mathcal{L}_{sup}$ and the individual pixel-level alignment $\mathcal{L}_{reg}$, RDC adds two components computed on pixel embeddings before the classifier: **(i)** *agent construction*, realized by orthogonal agent selection, which selects well-separated reference pixels (agents) from the teacher embedding and gathers them from both views with shared indices; and **(ii)** *order-based alignment supervision*, realized by a differentiable Kendall's τ, which aligns the order of the teacher and student pixel–agent affinities through $\mathcal{L}_{rdc}$. RDC is used only during training and introduces no cost at inference.

4 What the Student Learns from an Imperfect Teacher

The teacher's affinities vary with the weak augmentation and the EMA state, and each training step sees one realization of them. This section compares how value alignment and order alignment transmit this variation to the student's update.

Setup. For a pixel, write the teacher affinities as $\boldsymbol{z}^w = \boldsymbol{z}^* + \boldsymbol{\delta}$, where $\boldsymbol{z}^*$ is the teacher's mean affinity over weak views and EMA states and $\boldsymbol{\delta}$ is the deviation from it. For a relation (p, q), the reference gap is $u^*_{pq} = z^*_p - z^*_q$, the teacher gap is

$$u_{pq} = u^*_{pq} + \eta_{pq}, \qquad \eta_{pq} = \delta_p - \delta_q, \tag{11}$$

and the reference target is $\boldsymbol{c}^* = \mathrm{softmax}(\boldsymbol{z}^*)$. We take η_{pq} to be symmetric, $\eta_{pq} \overset{d}{=} -\eta_{pq}$, so u^*_{pq} is the median of u_{pq} and the reference direction $\mathrm{sign}(u^*_{pq})$ is the direction the teacher gives in at least half of the views; Proposition 1 does not require this symmetry. The student gap v_{pq} and the agent indices are fixed, and expectations are taken over $\boldsymbol{\delta}$.

Value alignment at the permutation level. The value alignment of Equation (5) compares affinity distributions. Aligning the rankings that these distributions generate contains the same comparison.

Lemma 1 (Permutation matching contains value matching)**.** *For a probability vector $\mathbf{c}$ over the N agents, let $P^{(k)}_{\mathbf{c}}$ be the Plackett–Luce distribution over the first k positions of a ranking, in which the agent at position t is drawn from $\mathbf{c}$ renormalized over the agents not yet placed, written $\boldsymbol{c}^{(t)}(\cdot \mid \pi_{1:t-1})$. For any $1 \le k \le N$ and probability vectors $\mathbf{c}, \mathbf{d}$,*

$$\mathrm{KL}\big(P^{(k)}_{\mathbf{c}} \,\|\, P^{(k)}_{\mathbf{d}}\big) = \mathrm{KL}(\boldsymbol{c} \,\|\, \boldsymbol{d}) + \sum_{t=2}^{k} \mathbb{E}_{\pi_{1:t-1} \sim P^{(t-1)}_{\mathbf{c}}} \Big[\mathrm{KL}\big(\boldsymbol{c}^{(t)}(\cdot \mid \pi_{1:t-1}) \,\|\, \boldsymbol{d}^{(t)}(\cdot \mid \pi_{1:t-1})\big)\Big]. \tag{12}$$

The affinity consistency of RankMatch (Mai et al., 2024) can be formalized as such a Plackett–Luce permutation KL between the teacher and student affinity distributions. In both its full ($k = N$) and top-k forms, it contains the value KL of Equation (5) as its first stage and adds value constraints on renormalized distributions at the later stages.

Order and value alignment under perturbation. The next result compares the two targets through the expected update that a relation, or an affinity distribution, induces in the student.

Theorem 1 (Order alignment keeps its direction; value alignment drifts)**.** *Fix a relation (p, q) with $u^*_{pq} \neq 0$, write $u = u^*_{pq} + \eta_{pq}$, and let ℓ be the relation loss of Equation (9).*

(a) ***Order.*** *Let $\kappa_{pq} = \mathbb{E}\, g_\alpha(u) \,/\, g_\alpha(u^*_{pq})$. For every student gap v,*

$$\mathbb{E}\,\ell(u, v) = 1 - \kappa_{pq} + \kappa_{pq}\,\ell(u^*_{pq}, v), \qquad \mathbb{E}\,\partial_v \ell(u, v) = \kappa_{pq}\,\partial_v \ell(u^*_{pq}, v). \tag{13}$$

The expected update keeps the reference direction if and only if $\kappa_{pq} > 0$, that is, if and only if $\mathbb{E}\big[w\,\mathbf{1}\{\operatorname{sign} u = \operatorname{sign} u^*_{pq}\}\big] > \mathbb{E}\big[w\,\mathbf{1}\{\operatorname{sign} u \neq \operatorname{sign} u^*_{pq}\}\big]$ with $w = |g_\alpha(u)|$; for $w \equiv 1$ this reduces to the classical condition that a label flips with probability below $1/2$ (Ghosh et al., 2017). Under the symmetry of the Setup, if the perturbation is non-degenerate, $\mathbb{P}(\eta_{pq} \neq 0) > 0$, then

$$\kappa_{pq} = \mathbb{E}\Big[\frac{1 - b^2}{1 - a^2 b^2}\Big] \in (0, 1), \qquad a = g_\alpha(u^*_{pq}), \quad b = g_\alpha(\eta_{pq}). \tag{14}$$

Consequently, if $\boldsymbol{z}^*$ has no ties and $\kappa_{pq} > 0$ for all relations of a pixel, the expected relational loss of the pixel equals a constant plus $\frac{1}{M}\sum_{p<q} \kappa_{pq}\,\ell(u^*_{pq}, v_{pq})$, with $M = \binom{N}{2}$: the perturbation reweights the relations, and every sequence of student affinities that approaches the infimum of this loss eventually orders the agents as $\boldsymbol{z}^*$ does.

(b) ***Values.*** *For any random teacher distribution $\boldsymbol{C}$ with mean $\bar{\boldsymbol{c}} = \mathbb{E}\boldsymbol{C}$ and any student distribution $\boldsymbol{d} = \operatorname{softmax}(\boldsymbol{z}^s)$,*

$$\mathbb{E}\,\mathrm{KL}(\boldsymbol{C} \,\|\, \boldsymbol{d}) = \mathbb{E}\,\mathrm{KL}(\boldsymbol{C} \,\|\, \bar{\boldsymbol{c}}) + \mathrm{KL}(\bar{\boldsymbol{c}} \,\|\, \boldsymbol{d}), \qquad \mathbb{E}\,\nabla_{\boldsymbol{z}^s}\mathrm{KL}(\boldsymbol{C} \,\|\, \boldsymbol{d}) = (\boldsymbol{d} - \boldsymbol{c}^*) + (\boldsymbol{c}^* - \bar{\boldsymbol{c}}), \tag{15}$$

where $\boldsymbol{c}^*$ is the reference target. For two agents with a non-degenerate perturbation, the target of value alignment shrinks toward $1/2$,

$$\bar{c}_p - \tfrac{1}{2} = \kappa^{(1)}_{pq}\big(c^*_p - \tfrac{1}{2}\big), \qquad 0 < \kappa^{(1)}_{pq} < 1, \tag{16}$$

where $\kappa^{(1)}_{pq}$ is given by Equation (14) with $\alpha = 1$. Hence, for every student gap v strictly between $\operatorname{logit}(\bar{c}_p)$ and u^*_{pq}, the expected update and the reference update have opposite signs.

Value alignment prescribes a distribution for the student to reach, and the perturbation moves this distribution from $\boldsymbol{c}^*$ to $\bar{\boldsymbol{c}}$. A student that lies between the two is then pushed away from $\boldsymbol{c}^*$, even when the perturbation never changes the order: with $u^* = 2$ and $\eta = \pm 1$, the teacher always prefers agent p, yet the target moves from $\operatorname{sigm}(2) \approx 0.881$ to 0.842. With more agents, the displacement can also reverse an order. The mean of a softmax favors the agents whose affinities vary more, because the exponential is convex: for independent $\delta_n \sim \mathcal{N}(0, s_n^2)$ and a normalizer that varies little, $\bar{c}_n$ is approximately proportional to $\exp(z^*_n + s_n^2/2)$, so the target of value alignment ranks the agents by $z^*_n + s_n^2/2$, and a noisier agent can overtake a higher one (Remark 1). Figure 5a traces this reversal, which occurs once $s > 0.68$. Value alignment thus converts zero-mean variation of the teacher into a systematic displacement, and since the teacher is an EMA of the student, the displacement absorbed by the student returns to the teacher. By Lemma 1, value alignment at the permutation level contains the value KL as its first stage, so its expected gradient carries the same drift term $\boldsymbol{c}^* - \bar{\boldsymbol{c}}$.

Order alignment prescribes only which agent should score higher. The perturbation multiplies its expected update by κ_{pq}, which is positive whenever the relation is preserved more often than inverted in the strength-weighted sense and lies in $(0, 1)$ under symmetric perturbations; each relation keeps its direction and changes only its strength. In the example of Figure 5b, the relation between the two agents is inverted with probability up to 0.46, while κ_{12} stays above 0.36. The strength a relation receives in expectation is $\kappa_{pq}\,|g_\alpha(u^*_{pq})|$, which is small for near-ties and under strong perturbations. The next result bounds how many relations are inverted.

Proposition 1 (Fraction of inverted relations)**.** *For a relation with $u^*_{pq} \neq 0$, let $\Delta_{pq} = |u^*_{pq}|$ and let $X_{pq} = -\operatorname{sign}(u^*_{pq})\,\eta_{pq}$, the part of the perturbation that pushes the relation toward inversion, have mean b_{pq}. If $X_{pq} - b_{pq}$ is ν_{pq}-sub-Gaussian, i.e. $\mathbb{E}\exp\{\lambda(X_{pq} - b_{pq})\} \leq \exp(\lambda^2\nu^2_{pq}/2)$ for all $\lambda \in \mathbb{R}$, then the relation is inverted with probability at most $\exp\{-(\Delta_{pq} - b_{pq})_+^2/(2\nu^2_{pq})\}$, where $(x)_+ = \max\{x, 0\}$, and the normalized Kendall disagreement $d_K(\boldsymbol{z}^w, \boldsymbol{z}^*) = \big(1 - \tau(\boldsymbol{z}^w, \boldsymbol{z}^*)\big)/2$ satisfies*

$$\mathbb{E}\,d_K(\boldsymbol{z}^w, \boldsymbol{z}^*) \leq \frac{1}{M}\sum_{p<q} \exp\Big(-\frac{(\Delta_{pq} - b_{pq})_+^2}{2\nu^2_{pq}}\Big). \tag{17}$$

The bound requires no independence across agents or relations. Applied to the teacher and the student, it controls their disagreement through $d_K(\boldsymbol{z}^w, \boldsymbol{z}^s) \le d_K(\boldsymbol{z}^w, \boldsymbol{z}^*) + d_K(\boldsymbol{z}^s, \boldsymbol{z}^*)$.

Equation (17) averages over all relations, as the Kendall loss does, so a few near-ties leave the supervision of the remaining relations intact. Averaging the exponential decay of individual relations over the margins of a pixel yields a power law in the noise scale, whose exponent is set by the fraction of normalized margins near zero (Remark 2). Proofs and remarks are given in Section B, and Section 5.4 measures these quantities on trained models.

Role of the agents. The margins Δ_{pq} depend on the agents. The orthogonal agent selection reduces the largest pairwise cosine $\rho(\boldsymbol{A}) = \max_{p \neq q} \hat{\boldsymbol{a}}_p^\top \hat{\boldsymbol{a}}_q$, with $\hat{\boldsymbol{a}}_n = \boldsymbol{a}_n / \|\boldsymbol{a}_n\|$. For agents of norm at least r, every pair satisfies $\|\boldsymbol{a}_p - \boldsymbol{a}_q\|^2 \ge 2r^2(1 - \rho(\boldsymbol{A}))$, and every pixel satisfies $\Delta_{pq} = |\boldsymbol{f}^\top(\boldsymbol{a}_p - \boldsymbol{a}_q)| \le \|\boldsymbol{f}\| \, \|\boldsymbol{a}_p - \boldsymbol{a}_q\|$. Near-duplicate agents therefore cap the margins of all pixels, and reducing $\rho(\boldsymbol{A})$ lifts this cap to at least $r\|\boldsymbol{f}\|\sqrt{2(1 - \rho(\boldsymbol{A}))}$. Section E.1 compares the selection with alternatives.

5 Experiments

Setup. We evaluate on four benchmarks: ACDC (Bernard et al., 2018) (2D cardiac MRI, four classes), LA (Xiong et al., 2021) (3D gadolinium-enhanced MRI), Pancreas-NIH (Roth et al., 2015) (3D CT), and Synapse (Chang et al., 2023) (abdominal CT, eight organs), following the splits and protocols of prior work (Bai et al., 2023; Chi et al., 2024). U-Net (Ronneberger et al., 2015) and V-Net (Milletari et al., 2016) serve as backbones for 2D and 3D data, and we report Dice, Jaccard, 95% Hausdorff distance (95HD), and average surface distance (ASD). RDC uses $N = 64$ agents, $\alpha = 2$, and $\lambda = 0.1$ on all datasets; Section A gives the remaining details. Baseline results are taken from the original papers under the same protocol, except those marked †, which we reproduce with the official code.

5.1 Comparison with State-of-the-art Methods

Tables 1–4 in Section A.1 compare RDC with semi-supervised segmentation methods. RDC achieves the best Dice in every setting and the lowest 95HD and ASD in all four ACDC and LA settings, reflecting more accurate boundaries. Over five identical seeds on ACDC-5% and LA-5%, it exceeds the two strongest baselines, GraphCL and GapMatch, by 1.37–2.10 Dice (paired t-test, $p < 0.01$; Table 11). VQ-Seg (Yang et al., 2026), which adds a vector-quantized reconstruction branch and foundation-model guidance, reaches 90.57 on ACDC-5% and 90.21 on LA-5%, and adding RDC raises them to 91.39 and 91.70 (Table 9). On Pancreas-NIH with 10% labels and on Synapse with one and two labeled cases, RDC improves the best baseline by 3.98, 4.0, and 2.4 Dice, with the largest relative gains on the organs whose pixels most often fall below the confidence threshold (Section E.2).

5.2 Order Alignment versus Value Alignment

Table 5 fixes the agents, the baseline, the pseudo-label loss, and λ, and changes only the alignment; CE is computed between the teacher and student affinity distributions. The five value alignments reach 88.68–89.03 Dice on ACDC-5% and the four order alignments reach 89.75–90.45, so every order alignment outperforms every value alignment. The Plackett–Luce KL, which formalizes the affinity consistency of RankMatch (Mai et al., 2024), performs like the other value alignments (88.71), as Lemma 1 anticipates. The advantage therefore comes from the choice of the alignment and holds across surrogates, and among the order alignments Kendall's τ, which directly counts the disagreeing relations bounded in Proposition 1, performs best. On ACDC, every target defined on agents also improves on contrastive learning over pixel pairs. Weighting each relation additionally by the teacher gap changes the Dice by 0.10, since the surrogate already weights each relation by w (Equation (9)); weighting it by the teacher's class confidence lowers the Dice to 89.71, as it suppresses the low-confidence pixels on which RDC contributes most (Section 5.3).

5.3 Complementarity with Pixel-level Pseudo-labels

The top of Table 6 isolates the two unlabeled terms. On its own, $\mathcal{L}_{rdc}$ raises the Dice of the supervised baseline by 30.08 and 26.48 on ACDC and LA, so order alignment carries substantial supervision. Combined with $\mathcal{L}_{reg}$, it adds 3.24 and 1.85, while $\mathcal{L}_{reg}$ adds 12.54 and 10.98 on top of $\mathcal{L}_{rdc}$: each term supplies information that the other lacks. The bottom of Table 6 varies the pixels that receive $\mathcal{L}_{rdc}$ according to the teacher's confidence. Restricting $\mathcal{L}_{rdc}$ to confident pixels recovers 1.94 of the 3.24 points, and the remaining 1.30 come from low-confidence pixels, which filtering leaves behind. Removing only the least confident decile changes the Dice by 0.14. The low-confidence pixels form thin bands along organ boundaries (Figure 11), and on them the inter-pixel supervision remains substantial: the mean relation strength $S_j = \frac{1}{M}\sum_{p<q} |g_\alpha(u_{j,pq})|$ of low-confidence pixels is 0.95 times that of high-confidence pixels.

5.4 Measurements of the Analysis

Table 7 measures the quantities of Section 4 on the trained ACDC-5% model. For Proposition 1, we apply the bound to the discrepancy between the student and the teacher, with the teacher affinities as reference, zero offset, and a single sub-Gaussian scale $\hat{\nu}^2 = 1.45$ whose envelope dominates the measured tail of $\delta_p - \delta_q$. The observed inversion rate stays below the bound in every stratum, and the bound is tightest in the largest-margin stratum, within 16% of the observed rate. We then test whether relations remain informative where the class decision fails. On labeled images, a relation (p, q) of pixel j whose agents satisfy $y_p = y_j \neq y_q$ is correct if $z_{jp} > z_{jq}$. The relation accuracy is 81.87% on pixels with a correct pseudo-label and 58.25% on pixels with an incorrect one, above the chance level of 50%: where the class decision fails, most relations still point the right way.

Proposition 1 also implies that near-ties carry little of the supervision. Removing every relation with $|\Delta_{pq}| < 0.1$ from $\mathcal{L}_{rdc}$ and retraining yields 90.49 Dice, on par with the 90.45 of RDC. Pixels with little dispersion among their affinities still carry well-separated relations: in the quintile of pixels with the least dispersed affinities, the mean margin is 1.12, or $0.93\hat{\nu}$, and these pixels contribute 15.5% of $\mathcal{L}_{rdc}$, against 24.2% for the most dispersed quintile.

Finally, Table 8 connects the analysis to segmentation accuracy. We inject noise into the teacher affinities used by the inter-pixel term and keep the pseudo-labels of $\mathcal{L}_{reg}$ unchanged, so that only the response of the alignment to the perturbation differs between KL and RDC. Theorem 1 predicts that the gap between the two alignments widens as the noise grows, which separates the effect of the perturbation from other differences between the losses, such as how they distribute gradients over agents (Guan et al., 2024). The gap grows from 1.47 without noise to 3.02 and 3.31 at the two noise levels: KL loses 2.19 and 3.56 Dice, whereas RDC loses 0.64 and 1.72. Section C reports the per-relation inversion probability, which decays exponentially in Δ^2/σ^2, the resulting aggregate scaling, and further measurements.

5.5 Generality and Cost

Table 9 adds RDC to three frameworks, improving GapMatch by 2.41 and 2.73 Dice on ACDC-5% and LA-5%, GraphCL by 2.34 and 2.07, and VQ-Seg by 0.82 and 1.49. In the SAM-based ESDE (Zhao et al., 2026), replacing the KL distillation of logits with the order loss improves LA-5% Dice by 1.89 (Section A.7). RDC adds 13.6% training time per iteration on LA-5% and none at inference (Table 13). The same $N = 64$, $\alpha = 2$, and $\lambda = 0.1$ are optimal on ACDC, LA, and Pancreas-NIH (Table 12); Section E.3 examines where order alignment helps less.

6 Conclusion

This paper studies how supervision from an imperfect teacher can be both reliable and rich in semi-supervised medical image segmentation. We propose RDC, which relaxes inter-pixel alignment from values to order. Noisy affinity values may still preserve the correct order, so learning the order provides reliable supervision on the uncertain pixels that filtering leaves behind. Our analysis shows that value alignment drifts under teacher noise and can reverse orders, whereas order alignment keeps the direction of each pairwise affinity order and only weakens it. Experiments on four benchmarks and four frameworks confirm the advantage of order alignment, and RDC adds no cost at inference.

AI USE STATEMENT

Generative AI tools were used solely for language editing and coding assistance. Specifically, they were used to improve the clarity, grammar, and presentation of the manuscript, as well as to assist with code implementation and debugging. They were not used to formulate the research problem, develop the core methodology, conduct experiments, analyze results, or draw scientific conclusions. All AI-assisted text was carefully reviewed and revised by the authors, and all AI-assisted code was inspected, tested, and validated for correctness. The authors take full responsibility for the final content of this work, including all text, claims, results, and code produced with the assistance of generative AI.

REFERENCES

Inigo Alonso, Alberto Sabater, David Ferstl, Luis Montesano, and Ana C Murillo. Semi-supervised semantic segmentation with pixel-level contrastive learning from a class-wise memory bank. In *Proceedings of the IEEE/CVF international conference on computer vision*, pp. 8219–8228, 2021.

Yunhao Bai, Duowen Chen, Qingli Li, Wei Shen, and Yan Wang. Bidirectional copy-paste for semi-supervised medical image segmentation. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 11514–11524, 2023.

Olivier Bernard, Alain Lalande, Clement Zotti, Frederick Cervenansky, Xin Yang, Pheng-Ann Heng, Irem Cetin, Karim Lekadir, Oscar Camara, Miguel Angel Gonzalez Ballester, et al. Deep learning techniques for automatic mri cardiac multi-structures segmentation and diagnosis: is the problem solved? *IEEE transactions on medical imaging*, 37(11):2514–2525, 2018.

Manuel Birlo, PJ Eddie Edwards, Matthew Clarkson, and Danail Stoyanov. Utility of optical see-through head mounted displays in augmented reality-assisted surgery: A systematic review. *Medical Image Analysis*, 77:102361, 2022.

Mathieu Blondel, Olivier Teboul, Quentin Berthet, and Josip Djolonga. Fast differentiable sorting and ranking. In *International conference on machine learning*, pp. 950–959. PMLR, 2020.

Chris Burges, Tal Shaked, Erin Renshaw, Ari Lazier, Matt Deeds, Nicole Hamilton, and Greg Hullender. Learning to rank using gradient descent. In *Proceedings of the 22nd international conference on Machine learning*, pp. 89–96, 2005.

Christopher Burges, Robert Ragno, and Quoc Le. Learning to rank with nonsmooth cost functions. *Advances in neural information processing systems*, 19, 2006.

Hu Cao, Yueyue Wang, Joy Chen, Dongsheng Jiang, Xiaopeng Zhang, Qi Tian, and Manning Wang. Swin-unet: Unet-like pure transformer for medical image segmentation. In *European conference on computer vision*, pp. 205–218. Springer, 2022.

Zhe Cao, Tao Qin, Tie-Yan Liu, Ming-Feng Tsai, and Hang Li. Learning to rank: from pairwise approach to listwise approach. In *Proceedings of the 24th international conference on Machine learning*, pp. 129–136, 2007.

Ao Chang, Xing Tao, Xin Yang, Yuhao Huang, Xinrui Zhou, Jiajun Zeng, Ruobing Huang, and Dong Ni. Pe-med: Prompt enhancement for interactive medical image segmentation. In *International workshop on machine learning in medical imaging*, pp. 257–266. Springer, 2023.

Hao Chen, Ran Tao, Yue Fan, Yidong Wang, Jindong Wang, Bernt Schiele, Xing Xie, Bhiksha Raj, and Marios Savvides. Softmatch: Addressing the quantity-quality trade-off in semi-supervised learning. *arXiv preprint arXiv:2301.10921*, 2023.

Xiaokang Chen, Yuhui Yuan, Gang Zeng, and Jingdong Wang. Semi-supervised semantic segmentation with cross pseudo supervision. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 2613–2622, 2021.

Hanyang Chi, Jian Pang, Bingfeng Zhang, and Weifeng Liu. Adaptive bidirectional displacement for semi-supervised medical image segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 4070–4080, 2024.

Aritra Ghosh, Himanshu Kumar, and P Shanti Sastry. Robust loss functions under label noise for deep neural networks. In *Proceedings of the AAAI conference on artificial intelligence*, volume 31, 2017.

Aditya Grover, Eric Wang, Aaron Zweig, and Stefano Ermon. Stochastic optimization of sorting networks via continuous relaxations. *arXiv preprint arXiv:1903.08850*, 2019.

Yuchen Guan, Runxi Cheng, Kang Liu, and Chun Yuan. Enhancing logits distillation with plug&play Kendall's τ ranking loss. *arXiv preprint arXiv:2409.17823*, 2024.

Chuan Guo, Geoff Pleiss, Yu Sun, and Kilian Q Weinberger. On calibration of modern neural networks. In *International conference on machine learning*, pp. 1321–1330. PMLR, 2017.

Tamás Haidegger. Autonomy for surgical robots: Concepts and paradigms. *IEEE Transactions on Medical Robotics and Bionics*, 1(2):65–76, 2019.

Huimin Huang, Yawen Huang, Shiao Xie, Lanfen Lin, Ruofeng Tong, Yen-Wei Chen, Yuexiang Li, and Yefeng Zheng. Combinatorial cnn-transformer learning with manifold constraints for semi-supervised medical image segmentation. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, pp. 2330–2338, 2024.

Wei Huang, Lei Zhang, Zizhou Wang, and Yan Wang. Gapmatch: Bridging instance and model perturbations for enhanced semi-supervised medical image segmentation. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, pp. 17458–17466, 2025.

Maurice G Kendall. A new measure of rank correlation. *Biometrika*, 30(1-2):81–93, 1938.

Alexander Kirillov, Eric Mintun, Nikhila Ravi, Hanzi Mao, Chloe Rolland, Laura Gustafson, Tete Xiao, Spencer Whitehead, Alexander C Berg, Wan-Yen Lo, et al. Segment anything. In *Proceedings of the IEEE/CVF international conference on computer vision*, pp. 4015–4026, 2023.

Tuomas Kynkäänniemi, Tero Karras, Miika Aittala, Timo Aila, and Jaakko Lehtinen. The role of imagenet classes in fr\'echet inception distance. *arXiv preprint arXiv:2203.06026*, 2022.

Tao Lei, Dong Zhang, Xiaogang Du, Xuan Wang, Yong Wan, and Asoke K Nandi. Semi-supervised medical image segmentation using adversarial consistency learning and dynamic convolution network. *IEEE transactions on medical imaging*, 42(5):1265–1277, 2022.

Shuailin Li, Chuyu Zhang, and Xuming He. Shape-aware semi-supervised 3d semantic segmentation for medical images. In *Medical Image Computing and Computer Assisted Intervention–MICCAI 2020: 23rd International Conference, Lima, Peru, October 4–8, 2020, Proceedings, Part I 23*, pp. 552–561. Springer, 2020.

Shumeng Li, Lei Qi, Qian Yu, Jing Huo, Yinghuan Shi, and Yang Gao. Stitching, fine-tuning, and re-training: A sam-enabled framework for semi-supervised 3d medical image segmentation. *IEEE Transactions on Medical Imaging*, 44(10):3909–3923, 2025.

Yuyuan Liu, Yu Tian, Yuanhong Chen, Fengbei Liu, Vasileios Belagiannis, and Gustavo Carneiro. Perturbed and strict mean teachers for semi-supervised semantic segmentation. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 4258–4267, 2022.

Wenjing Lu, Yi Hong, and Yang Yang. Up-sam: Uncertainty-informed adaptation of segment anything model for semi-supervised medical image segmentation. In *2024 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, pp. 2256–2261. IEEE, 2024.

Xiangde Luo, Jieneng Chen, Tao Song, and Guotai Wang. Semi-supervised medical image segmentation through dual-task consistency. In *Proceedings of the AAAI conference on artificial intelligence*, volume 35, pp. 8801–8809, 2021a.

Xiangde Luo, Wenjun Liao, Jieneng Chen, Tao Song, Yinan Chen, Shichuan Zhang, Nianyong Chen, Guotai Wang, and Shaoting Zhang. Efficient semi-supervised gross target volume of nasopharyngeal carcinoma segmentation via uncertainty rectified pyramid consistency. In *Medical Image Computing and Computer Assisted Intervention–MICCAI 2021: 24th International Conference, Strasbourg, France, September 27–October 1, 2021, Proceedings, Part II 24*, pp. 318–329. Springer, 2021b.

Jun Ma, Yuting He, Feifei Li, Lin Han, Chenyu You, and Bo Wang. Segment anything in medical images. *Nature communications*, 15(1):654, 2024.

Xingjun Ma, Hanxun Huang, Yisen Wang, Simone Romano, Sarah Erfani, and James Bailey. Normalized loss functions for deep learning with noisy labels. In *International conference on machine learning*, pp. 6543–6553. PMLR, 2020.

Huayu Mai, Rui Sun, Tianzhu Zhang, and Feng Wu. Rankmatch: Exploring the better consistency regularization for semi-supervised semantic segmentation. In *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 3391–3401. IEEE, 2024.

Juzheng Miao, Cheng Chen, Furui Liu, Hao Wei, and Pheng-Ann Heng. Caussl: Causality-inspired semi-supervised learning for medical image segmentation. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 21426–21437, 2023.

Fausto Milletari, Nassir Navab, and Seyed-Ahmad Ahmadi. V-net: Fully convolutional neural networks for volumetric medical image segmentation. In *2016 fourth international conference on 3D vision (3DV)*, pp. 565–571. Ieee, 2016.

Giorgio Patrini, Alessandro Rozza, Aditya Krishna Menon, Richard Nock, and Lizhen Qu. Making deep neural networks robust to label noise: A loss correction approach. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1944–1952, 2017.

Sebastian Prillo and Julian Eisenschlos. Softsort: A continuous relaxation for the argsort operator. In *International Conference on Machine Learning*, pp. 7793–7802. PMLR, 2020.

Babak Rezaeirowshan, Coloma Ballester, and Gloria Haro. Monocular depth ordering using perceptual occlusion cues. In *Proceedings of the 11th Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*, 2016.

Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18*, pp. 234–241. Springer, 2015.

Holger R Roth, Le Lu, Amal Farag, Hoo-Chang Shin, Jiamin Liu, Evrim B Turkbey, and Ronald M Summers. Deeporgan: Multi-level deep convolutional networks for automated pancreas segmentation. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part I 18*, pp. 556–564. Springer, 2015.

Florian Schroff, Dmitry Kalenichenko, and James Philbin. Facenet: A unified embedding for face recognition and clustering. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 815–823, 2015.

Kihyuk Sohn, David Berthelot, Nicholas Carlini, Zizhao Zhang, Han Zhang, Colin A Raffel, Ekin Dogus Cubuk, Alexey Kurakin, and Chun-Liang Li. Fixmatch: Simplifying semi-supervised learning with consistency and confidence. *Advances in neural information processing systems*, 33:596–608, 2020.

Charles Spearman. The proof and measurement of association between two things. 1961.

Antti Tarvainen and Harri Valpola. Mean teachers are better role models: Weight-averaged consistency targets improve semi-supervised deep learning results. *Advances in neural information processing systems*, 30, 2017.

Haoyu Wang, Sizheng Guo, Jin Ye, Zhongying Deng, Junlong Cheng, Tianbin Li, Jianpin Chen, Yanzhou Su, Ziyan Huang, Yiqing Shen, et al. Sam-med3d: a vision foundation model for general-purpose segmentation on volumetric medical images. *IEEE Transactions on Neural Networks and Learning Systems*, 2025.

Mengzhu Wang, Jiao Li, Houcheng Su, Nan Yin, Liang Yang, and Shen Li. Graphcl: Graph-based clustering for semi-supervised medical image segmentation. *arXiv preprint arXiv:2411.13147*, 2024.

Wenguan Wang, Tianfei Zhou, Fisher Yu, Jifeng Dai, Ender Konukoglu, and Luc Van Gool. Exploring cross-image pixel contrast for semantic segmentation. In *Proceedings of the IEEE/CVF international conference on computer vision*, pp. 7303–7313, 2021.

Yidong Wang, Hao Chen, Qiang Heng, Wenxin Hou, Marios Savvides, Takahiro Shinozaki, Bhiksha Raj, Zhen Wu, and Jindong Wang. Freematch: Self-adaptive thresholding for semi-supervised learning. *arXiv preprint arXiv:2205.07246*, 2022.

Yisen Wang, Xingjun Ma, Zaiyi Chen, Yuan Luo, Jinfeng Yi, and James Bailey. Symmetric cross entropy for robust learning with noisy labels. In *Proceedings of the IEEE/CVF international conference on computer vision*, pp. 322–330, 2019.

Yongchao Wang, Bin Xiao, Xiuli Bi, Weisheng Li, and Xinbo Gao. Mcf: Mutual correction framework for semi-supervised medical image segmentation. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 15651–15660, 2023.

Colin Wei, Kendrick Shen, Yining Chen, and Tengyu Ma. Theoretical analysis of self-training with deep networks on unlabeled data. *arXiv preprint arXiv:2010.03622*, 2020.

Junde Wu, Ziyue Wang, Mingxuan Hong, Wei Ji, Huazhu Fu, Yanwu Xu, Min Xu, and Yueming Jin. Medical sam adapter: Adapting segment anything model for medical image segmentation. *Medical image analysis*, 102:103547, 2025.

Yicheng Wu, Minfeng Xu, Zongyuan Ge, Jianfei Cai, and Lei Zhang. Semi-supervised left atrium segmentation with mutual consistency training. In *Medical image computing and computer assisted intervention–MICCAI 2021: 24th international conference, Strasbourg, France, September 27–October 1, 2021, proceedings, part II 24*, pp. 297–306. Springer, 2021.

Yicheng Wu, Zhonghua Wu, Qianyi Wu, Zongyuan Ge, and Jianfei Cai. Exploring smoothness and class-separation for semi-supervised medical image segmentation. In *International conference on medical image computing and computer-assisted intervention*, pp. 34–43. Springer, 2022.

Fen Xia, Tie-Yan Liu, Jue Wang, Wensheng Zhang, and Hang Li. Listwise approach to learning to rank: theory and algorithm. In *Proceedings of the 25th international conference on Machine learning*, pp. 1192–1199, 2008.

Zhaohan Xiong, Qing Xia, Zhiqiang Hu, Ning Huang, Cheng Bian, Yefeng Zheng, Sulaiman Vesal, Nishant Ravikumar, Andreas Maier, Xin Yang, et al. A global benchmark of algorithms for segmenting the left atrium from late gadolinium-enhanced cardiac magnetic resonance imaging. *Medical image analysis*, 67:101832, 2021.

Lihe Yang, Lei Qi, Litong Feng, Wayne Zhang, and Yinghuan Shi. Revisiting weak-to-strong consistency in semi-supervised semantic segmentation. *arXiv preprint arXiv:2208.09910*, 2022.

Sicheng Yang, Zhaohu Xing, and Lei Zhu. Vq-seg: Vector-quantized token perturbation for semi-supervised medical image segmentation. *Advances in Neural Information Processing Systems*, 38: 22148–22172, 2026.

Chenyu You, Ruihan Zhao, Lawrence H Staib, and James S Duncan. Momentum contrastive voxel-wise representation learning for semi-supervised volumetric medical image segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 639–652. Springer, 2022.

Chenyu You, Weicheng Dai, Yifei Min, Fenglin Liu, David Clifton, S Kevin Zhou, Lawrence Staib, and James Duncan. Rethinking semi-supervised medical image segmentation: A variance-reduction perspective. *Advances in neural information processing systems*, 36:9984–10021, 2023.

Lequan Yu, Shujun Wang, Xiaomeng Li, Chi-Wing Fu, and Pheng-Ann Heng. Uncertainty-aware self-ensembling model for semi-supervised 3d left atrium segmentation. In *Medical image computing and computer assisted intervention–MICCAI 2019: 22nd international conference, Shenzhen, China, October 13–17, 2019, proceedings, part II 22*, pp. 605–613. Springer, 2019.

Bowen Zhang, Yidong Wang, Wenxin Hou, Hao Wu, Jindong Wang, Manabu Okumura, and Takahiro Shinozaki. Flexmatch: Boosting semi-supervised learning with curriculum pseudo labeling. *Advances in Neural Information Processing Systems*, 34:18408–18419, 2021.

Kaidong Zhang and Dong Liu. Customized segment anything model for medical image segmentation. *arXiv preprint arXiv:2304.13785*, 2023.

Yichi Zhang, Jin Yang, Yuchen Liu, Yuan Cheng, and Yuan Qi. Semisam: Enhancing semi-supervised medical image segmentation via sam-assisted consistency regularization. In *2024 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, pp. 3982–3986. IEEE, 2024.

Yichi Zhang, Bohao Lv, Le Xue, Wenbo Zhang, Yuchen Liu, Yu Fu, Yuan Cheng, and Yuan Qi. Semisam+: rethinking semi-supervised medical image segmentation in the era of foundation models. *Medical Image Analysis*, pp. 103733, 2025.

Zhilu Zhang and Mert Sabuncu. Generalized cross entropy loss for training deep neural networks with noisy labels. *Advances in neural information processing systems*, 31, 2018.

Haifeng Zhao, Haiyang Li, Lei-Lei Ma, and Dengdi Sun. Segment anything model meets semi-supervised medical image segmentation: A novel perspective. *Advances in Neural Information Processing Systems*, 38:170128–170148, 2026.

Xinkai Zhao, Chaowei Fang, De-Jun Fan, Xutao Lin, Feng Gao, and Guanbin Li. Cross-level contrastive learning and consistency constraint for semi-supervised medical image segmentation. In *2022 IEEE 19th international symposium on biomedical imaging (ISBI)*, pp. 1–5. IEEE, 2022.

Kaipeng Zheng, Huishuai Zhang, and Weiran Huang. Diffkendall: a novel approach for few-shot learning with differentiable kendall's rank correlation. *Advances in Neural Information Processing Systems*, 36:49403–49415, 2023.

Jiayuan Zhu, Abdullah Hamdi, Yunli Qi, Yueming Jin, and Junde Wu. Medical sam 2: Segment medical images as video via segment anything model 2. *arXiv preprint arXiv:2408.00874*, 2024.

A Experimental Details and Additional Results

A.1 Result Tables of Section 5

On ACDC and LA, the compared methods are UA-MT (Yu et al., 2019), SASSNet (Li et al., 2020), DTC (Luo et al., 2021a), URPC (Luo et al., 2021b), SS-Net (Wu et al., 2022), MC-Net+ (Wu et al., 2021), PS-MT (Liu et al., 2022), MCF (Wang et al., 2023), CauSSL (Miao et al., 2023), ABD (Chi et al., 2024), M-CnT (Huang et al., 2024), BCP (Bai et al., 2023), GapMatch (Huang et al., 2025), GraphCL (Wang et al., 2024), ARCO (You et al., 2023), and the SAM-based SemiSAM (Zhang et al., 2024) and UP-SAM (Lu et al., 2024). On Pancreas-NIH, they are UA-MT, DTC, ASE-Net (Lei et al., 2022), SS-Net, PS-MT, BCP, and CauSSL; on Synapse, DTC, MC-Net+, BCP, and ABD.

Table 1: Comparison on **ACDC** under 5% and 10% labeled data ($m\boldsymbol{x}^l$ + $n\boldsymbol{x}^u$: m labeled and n unlabeled scans). †: reproduced under our protocol. Best in **bold**.

Method	5% / $3\boldsymbol{x}^l$ + $67\boldsymbol{x}^u$				10% / $7\boldsymbol{x}^l$ + $63\boldsymbol{x}^u$			
	Dice ↑	Jaccard ↑	95HD ↓	ASD ↓	Dice ↑	Jaccard ↑	95HD ↓	ASD ↓
U-Net (*Sup.Only*)	47.83	37.01	31.16	12.62	79.41	68.11	9.35	2.70
UA-MT	46.04	35.97	20.08	7.75	81.65	70.64	6.88	2.02
DTC	56.90	45.67	23.36	7.39	84.29	73.92	12.81	4.01
URPC	55.87	44.64	13.60	3.74	83.10	72.41	4.84	1.53
SS-Net	65.82	55.38	6.67	2.28	86.78	77.67	6.07	1.40
MC-Net+	62.85	52.29	7.62	2.33	87.10	76.06	6.68	2.00
PS-MT	86.94	77.90	4.95	2.18	88.91	80.79	4.96	1.83
CauSSL	-	-	-	-	89.66	81.79	3.67	0.93
ABD	88.96	80.70	1.57	0.52	89.81	81.95	1.46	0.49
M-CnT	-	-	-	-	88.40	-	4.40	-
BCP	87.59	78.67	1.90	0.67	88.84	80.62	3.98	1.17
GapMatch	88.80	-	-	0.64	90.06	-	-	0.52
GraphCL	88.68	80.32	13.12	0.88	89.31	81.33	2.10	0.66
ARCO†	88.53	-	-	-	-	-	-	-
RDC (Ours)	**90.45**	**82.07**	**1.25**	**0.41**	**91.00**	**82.51**	**1.28**	**0.39**

Table 2: Comparison on **LA** under 5% and 10% labeled data. Best in **bold**.

Method	5% / $4\boldsymbol{x}^l$ + $76\boldsymbol{x}^u$				10% / $8\boldsymbol{x}^l$ + $72\boldsymbol{x}^u$			
	Dice ↑	Jaccard ↑	95HD ↓	ASD ↓	Dice ↑	Jaccard ↑	95HD ↓	ASD ↓
V-Net (*Sup.Only*)	52.55	39.60	47.05	9.87	82.74	71.72	13.35	3.26
UA-MT	82.26	70.98	13.71	3.82	86.28	76.11	18.71	4.63
SASSNet	81.60	69.63	16.16	3.58	85.22	75.09	11.18	2.89
DTC	81.25	69.33	14.90	3.99	87.51	78.17	8.23	2.36
URPC	82.48	71.35	14.65	3.65	85.01	74.36	15.37	3.96
SS-Net	86.33	76.15	9.97	2.31	88.55	79.62	7.49	1.90
MC-Net+	83.59	72.36	14.07	2.70	88.96	80.25	7.93	1.86
PS-MT	88.49	79.13	8.12	2.78	89.72	81.48	6.94	1.92
MCF	-	-	-	-	88.71	80.41	6.32	1.90
BCP	88.02	78.72	7.90	2.15	89.62	81.31	6.81	1.76
SemiSAM	80.42	68.05	18.23	5.16	84.45	73.75	14.56	3.31
GapMatch	88.30	-	-	1.92	90.13	-	-	1.82
UP-SAM	84.06	72.90	13.78	3.14	-	-	-	-
GraphCL	88.80	80.00	7.16	2.10	90.24	82.31	6.42	1.71
RDC (Ours)	**90.01**	**82.14**	**6.07**	**1.58**	**90.98**	**83.27**	**5.20**	**1.42**

Table 3: Comparison on **Pancreas-NIH** (full results in Table 10).

Method	10% labeled		20% labeled	
	Dice ↑	95HD ↓	Dice ↑	95HD ↓
V-Net (*Sup.Only*)	55.60	45.33	72.38	19.35
UA-MT	66.34	17.21	76.10	10.84
DTC	69.21	17.21	78.27	8.36
ASE-Net	71.54	13.33	79.03	8.62
SS-Net	71.76	17.56	79.78	8.86
PS-MT	76.94	13.12	80.74	7.41
BCP	73.83	12.71	82.91	6.43
CauSSL	72.89	14.19	80.92	8.11
RDC (Ours)	**80.92**	**6.53**	**83.37**	**4.28**

Table 4: Comparison on **Synapse** (Dice %).

Method	Labeled cases	
	1	2
Sup.-only	10.7	42.5
DTC	15.0	48.8
MC-Net+	26.3	55.2
BCP	34.0	61.1
ABD	41.1	64.0
RDC (Ours)	**45.1**	**66.4**

Table 5: Alignments added to the same baseline with the same agents and λ (Dice % on ACDC-5%).

Alignment	Dice
None ($\mathcal{L}_{sup} + \mathcal{L}_{reg}$)	87.21
Contrastive (SS-Net)†	88.25
Value alignment	
L2	88.68
CE	88.79
KL (DC)	88.98
Plackett–Luce KL	88.71
SoftSort	89.03
Order alignment	
RankNet	89.75
Pairwise hinge	90.09
Spearman	90.23
RDC	**90.45**

Table 6: Unlabeled terms (top) and pixels receiving $\mathcal{L}_{rdc}$ according to teacher confidence (bottom), Dice %.

$\mathcal{L}_{sup}$	$\mathcal{L}_{reg}$	$\mathcal{L}_{rdc}$	ACDC	LA
✓			47.83	52.55
✓		✓	77.91	79.03
✓	✓		87.21	88.16
✓	✓	✓	**90.45**	**90.01**

Pixels receiving $\mathcal{L}_{rdc}$	ACDC
None	87.21
High-confidence only	89.15
Low-confidence only	87.99
Least confident decile only	87.30
All but the least confident decile	90.31
All pixels (RDC)	**90.45**

Table 7: Measurements of Section 4 on the trained ACDC-5% model. (a) Inversion rate of relations between student and teacher, stratified by the reference margin Δ in units of the sub-Gaussian scale $\hat{\nu}$, and the bound of Proposition 1. (b) Relation accuracy on labeled images. (c) Training without near-tied relations.

(a) Proposition 1				
$\Delta/\hat{\nu}$	$[0, 0.5)$	$[0.5, 1)$	$[1, 2)$	$[2, 3)$
Observed	0.343	0.171	0.098	0.057
Bound	0.962	0.761	0.388	0.066
(b) Relation accuracy				
Pixels with a correct pseudo-label	81.87%			
Pixels with an incorrect pseudo-label	58.25%			
(c) Near-ties				
Dice without relations with $\lvert\Delta_{pq}\rvert < 0.1$	90.49 (RDC: 90.45)			

Table 8: Dice (%) on ACDC-5% with noise on the teacher affinities of the inter-pixel term only (pseudo-labels unchanged).

Alignment	None	Low	High
KL (value)	88.98	86.79	85.42
RDC (order)	**90.45**	**89.81**	**88.73**
Gap	1.47	3.02	3.31

Table 9: RDC as a plug-in (Dice %).

Method	ACDC	LA
GapMatch (Huang et al., 2025)	88.80	88.30
+ RDC	**91.21**	**91.03**
GraphCL (Wang et al., 2024)	88.68	88.80
+ RDC	**91.02**	**90.87**
VQ-Seg (Yang et al., 2026)	90.57	90.21
+ RDC	**91.39**	**91.70**

A.2 Setup

Datasets. (1) **ACDC** (Bernard et al., 2018) provides 2D cardiac cine-MRI with four classes from 100 patients (70/10/20 split). (2) **LA** (Xiong et al., 2021) comprises 100 3D GE-MRI scans (80/20 split). (3) **Pancreas-NIH** (Roth et al., 2015) contains 82 abdominal CT volumes (62/20 split). (4) **Synapse** (Chang et al., 2023) contains 30 axial abdominal CT scans (3,779 slices) with 8 annotated organs (aorta, gallbladder, left and right kidneys, liver, pancreas, spleen, and stomach) and strong class imbalance, the four smallest organs accounting for less than 3.5% of the foreground pixels. We follow prior work for the split and evaluate with one and two labeled cases using U-Net.

Implementation details. Following Yu et al. (2019); Bai et al. (2023); Chi et al. (2024), we use U-Net (Ronneberger et al., 2015) for 2D data (ACDC, Synapse) and V-Net (Milletari et al., 2016) for 3D data (LA, Pancreas-NIH). For ACDC, we use batch size 24, crop size 256×256, and 30K iterations. For LA and Pancreas-NIH, we use batch size 4, 15K iterations, and crop sizes $96 \times 96 \times 96$ and $112 \times 112 \times 80$. The student is trained with SGD (polynomial decay, learning rate 0.01), with EMA momentum $m = 0.99$ and confidence threshold $\gamma = 0.95$ (2D) and 0.75 (3D). The strong augmentation includes color jitter, Gaussian blur, and CutMix, and the weak augmentation uses random rotation and flip. RDC uses $N = 64$ agents, temperature $\alpha = 2$, and weight $\lambda = 0.1$ on all datasets. All experiments run on one RTX 3090.

A.3 Full Results and Qualitative Comparisons

Table 10 reports all four metrics on Pancreas-NIH. Figure 3 shows Synapse cases: in the top row, BCP and ABD predict stomach where none exists, and RDC avoids this false positive; in the bottom row, RDC segments the pancreas, a difficult organ with variable shape and low contrast, with minor boundary errors under this extreme label scarcity. Figure 4 compares predictions on ACDC and LA with 5% labels. ARCO (You et al., 2023) uses a 120/40/40 split of ACDC; we re-run it under our 70/10/20 split with its public code and obtain 88.53 Dice with 5% labels, close to the independent reproduction of 88.79 reported by VQ-Seg (Yang et al., 2026).

Table 10: Comparison on **Pancreas-NIH** under 10% and 20% labeled data.

Method	10% / $6\boldsymbol{x}^l$ + $56\boldsymbol{x}^u$				20% / $12\boldsymbol{x}^l$ + $50\boldsymbol{x}^u$			
	Dice ↑	Jaccard ↑	95HD ↓	ASD ↓	Dice ↑	Jaccard ↑	95HD ↓	ASD ↓
V-Net (*Sup.Only*)	55.60	41.74	45.33	18.63	72.38	58.26	19.35	5.89
UA-MT	66.34	53.21	17.21	4.57	76.10	62.62	10.84	2.43
SASSNet	68.78	53.86	19.02	6.26	77.66	64.08	10.93	3.05
DTC	69.21	54.06	17.21	5.95	78.27	64.75	8.36	2.25
ASE-Net	71.54	56.82	13.33	5.73	79.03	66.57	8.62	2.30
SS-Net	71.76	57.05	17.56	5.77	79.78	66.32	8.86	2.01
MC-Net+	70.00	55.66	16.03	3.87	79.37	66.83	8.52	1.72
PS-MT	76.94	62.37	13.12	3.66	80.74	68.15	7.41	2.06
MCF	-	-	-	-	75.00	61.27	11.59	3.27
BCP	73.83	59.24	12.71	3.72	82.91	70.97	6.43	2.25
CauSSL	72.89	58.06	14.19	1.66	80.92	68.26	8.11	1.53
RDC (Ours)	**80.92**	**68.24**	**6.53**	**1.41**	**83.37**	**71.74**	**4.28**	**1.33**

A.4 Repeated Runs

Table 11 compares RDC with the two strongest baselines, which we retrain with the same five seeds. On both datasets, the worst run of RDC is better than the best run of either baseline, and every paired t-test is significant at $p < 0.01$.

A.5 Hyperparameters

Table 12 varies the number of agents N, the temperature α, and the weight λ on three datasets. Each hyperparameter has the same optimum on all three. Few agents describe a pixel coarsely, while many

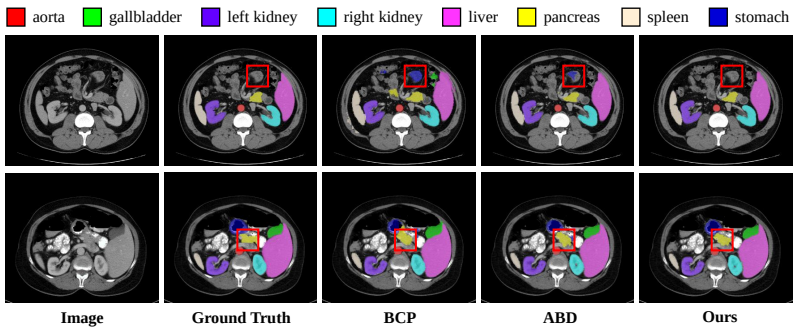


Figure 3: Qualitative comparison on Synapse. *Top*: RDC avoids a false positive of BCP and ABD. *Bottom*: the pancreas under extreme label scarcity.

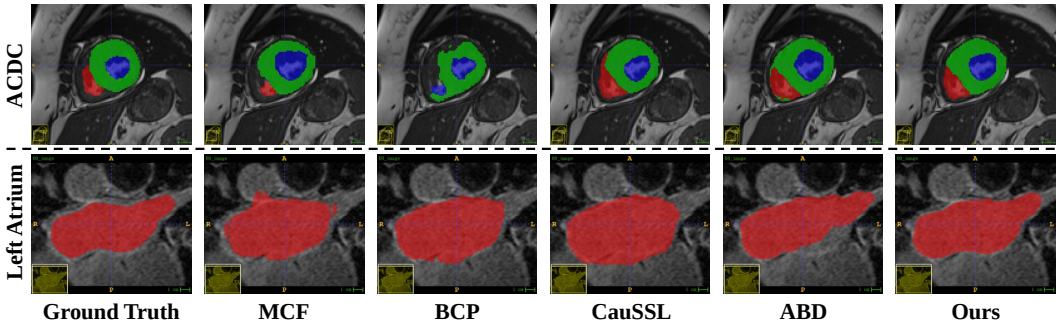


Figure 4: Qualitative comparison on ACDC (5%) and LA (5%).

agents crowd the affinities and shrink the margins (Section 4); a small α leaves the surrogate far from Kendall's τ, while a large α saturates g_α and weakens the gradient.

A.6 Training Cost

Table 13 breaks down the training time on LA-5%. Agent selection and the order loss add $0.22\,\mathrm{s}$ per iteration (13.6%), computed as batched matrix products over the $\binom{64}{2} = 2016$ relations of each voxel, and $0.4\,\mathrm{GB}$ of memory. Agents, affinities, and the loss are removed at test time, so inference is unchanged.

A.7 Compatibility with Foundation Models

We integrate $\mathcal{L}_{rdc}$ into ESDE (Zhao et al., 2026), a SAM-based SSMIS framework that distills a SAM teacher fine-tuned with LoRA into a TinyViT-5M student, with a KL divergence on softmax logits (their $\mathcal{L}_{\text{logit}}$) as the distillation loss. We replace this value alignment with the order loss and keep the rest of ESDE (embedding-level distillation, dynamic loss weighting, and two-stage fine-tuning) unchanged. On LA-5% (Table 14), the replacement improves ESDE by 1.89 Dice and 1.73 Jaccard and reduces the 95HD from 8.15 to 7.69 and the ASD from 2.65 to 2.54.

A.8 Quality of Pseudo-labels

We measure the Fréchet Inception Distance (FID) (Kynkäänniemi et al., 2022) between pseudo-labels and ground truth on LA (Table 15). RDC lowers the FID from 3.41 to 1.84 with 5% labels and from 2.73 to 1.53 with 10% labels, so its pseudo-labels are closer to the ground-truth label distribution.

Table 11: Five runs with identical seeds (Dice %). Δ: mean improvement of RDC; paired t-test with four degrees of freedom. †: reproduced.

Setting	Method	Mean ± std	Δ	t_4	p
ACDC 5%	GraphCL† (Wang et al., 2024)	88.68 ± 0.76	$+1.77$	5.35	0.006
	GapMatch† (Huang et al., 2025)	88.65 ± 0.46	$+1.80$	8.35	0.001
	RDC	$\mathbf{90.45 \pm 0.15}$	–	–	–
LA 5%	GraphCL† (Wang et al., 2024)	88.64 ± 0.51	$+1.37$	8.48	0.001
	GapMatch† (Huang et al., 2025)	87.91 ± 0.74	$+2.10$	5.40	0.006
	RDC	$\mathbf{90.01 \pm 0.19}$	–	–	–

Table 12: Sensitivity to N, α, and λ (Dice %). The deployed setting is shaded.

	Value	ACDC 5%	LA 5%	Pancreas-NIH 10%
N	32	89.55	88.70	79.87
	64	**90.45**	**90.01**	**80.92**
	128	89.89	89.47	80.13
	256	89.37	88.21	79.70
α	1	89.99	89.56	80.51
	2	**90.45**	**90.01**	**80.92**
	3	89.31	89.19	80.03
λ	0.05	89.51	89.18	79.97
	0.1	**90.45**	**90.01**	**80.92**
	0.2	89.90	89.35	80.14
	0.5	89.31	88.71	79.28

Table 13: Training time per iteration on LA-5% (V-Net, batch 4, one RTX 3090, forward and backward, mean ± std over three runs).

Component	Time per iteration (s)
Baseline (forward, $\mathcal{L}_{sup}$, $\mathcal{L}_{reg}$, backward)	1.62 ± 0.02
+ agent selection	$+0.08 \pm 0.01$
+ affinities, $\mathcal{L}_{rdc}$, and its backward pass	$+0.14 \pm 0.01$
Total with RDC	1.84 ± 0.03

Table 14: SAM-based SSMIS frameworks on **LA** (5% labeled). Best in **bold**.

Method	Dice ↑	Jaccard ↑	95HD ↓	ASD ↓
SemiSAM (Zhang et al., 2024)	80.42	68.05	18.23	5.16
UP-SAM (Lu et al., 2024)	84.06	72.90	13.78	3.14
ESDE (Zhao et al., 2026)	88.42	79.47	8.15	2.65
ESDE + RDC (Ours)	**90.31**	**81.20**	**7.69**	**2.54**

Table 15: FID (↓) between pseudo-labels and ground truth on LA.

Method	5%	10%
Baseline	3.41	2.73
RDC	**1.84**	**1.53**

B Proofs and Remarks for Section 4

Proof of Lemma 1. The Plackett–Luce distribution factorizes as $P_{\boldsymbol{c}}^{(k)}(\pi_{1:k}) = \prod_{t=1}^{k} \boldsymbol{c}^{(t)}(\pi_t \mid \pi_{1:t-1})$ with $\boldsymbol{c}^{(1)} = \boldsymbol{c}$, and likewise for $\boldsymbol{d}$. The chain rule of relative entropy applied to this factorization gives Equation (12). □

Proof of Theorem 1(a). Since $\ell(u,v) = 1 - g_\alpha(u)g_\alpha(v)$ is affine in $g_\alpha(u)$, $\mathbb{E}\,\ell(u,v) = 1 - \mathbb{E}[g_\alpha(u)]\,g_\alpha(v) = 1 - \kappa_{pq}\,a\,g_\alpha(v)$ with $a = g_\alpha(u^*_{pq})$, which is the first identity in Equation (13); differentiating in v gives the second. Writing $g_\alpha(u) = \operatorname{sign}(u)\,w$ gives

$$\operatorname{sign}(u^*_{pq})\,\mathbb{E}\,g_\alpha(u) = \mathbb{E}\big[w\,\mathbf{1}\{\operatorname{sign} u = \operatorname{sign} u^*_{pq}\}\big] - \mathbb{E}\big[w\,\mathbf{1}\{\operatorname{sign} u \neq \operatorname{sign} u^*_{pq}\}\big],$$

and $\kappa_{pq} > 0$ exactly when this quantity is positive. For the symmetric case, write $A = \alpha u^*_{pq}/2$ and $B = \alpha\eta_{pq}/2$, so that $a = \tanh A$ and $b = \tanh B$. The addition formula for the hyperbolic tangent gives

$$\frac{\tanh(A+B) + \tanh(A-B)}{2} = \frac{a(1-b^2)}{1-a^2b^2}.$$

Since η_{pq} and $-\eta_{pq}$ have the same distribution, $\mathbb{E}\,g_\alpha(u)$ equals the expectation of the left-hand side, hence $\kappa_{pq} = \mathbb{E}[(1-b^2)/(1-a^2b^2)]$. For finite arguments, $|a| < 1$ and $|b| < 1$, so the integrand lies in $(0,1]$ and is strictly smaller than one whenever $b \neq 0$; non-degeneracy gives $0 < \kappa_{pq} < 1$.

For the pixel-level statement, the loss of a pixel is $1 - \hat{\tau}_\alpha = \frac{1}{M}\sum_{p<q} \ell(u_{pq}, v_{pq})$, so its expectation is $\frac{1}{M}\sum_{p<q}(1-\kappa_{pq}) + \frac{1}{M}\sum_{p<q}\kappa_{pq}\,\ell(u^*_{pq}, v_{pq})$. Each term $\kappa_{pq}\,\ell(u^*_{pq}, v_{pq}) = \kappa_{pq}\big(1 - a_{pq}\,g_\alpha(v_{pq})\big)$, with $a_{pq} = g_\alpha(u^*_{pq})$, decreases strictly as $\operatorname{sign}(a_{pq})\,v_{pq}$ grows and tends to $\kappa_{pq}(1-|a_{pq}|)$. The student affinities $\boldsymbol{z}^s = t\,\boldsymbol{z}^*$ with $t \to \infty$ attain all these limits at once, so the infimum of the expected loss is the sum of the limits. A student that disagrees with $\boldsymbol{z}^*$ on some relation keeps that term at least κ_{pq}, which exceeds its limit by $\kappa_{pq}|a_{pq}| > 0$; hence a sequence approaching the infimum can disagree on a relation only finitely often. □

Proof of Theorem 1(b). Expanding $\mathbb{E}\sum_i C_i \log(C_i/d_i) = \mathbb{E}\sum_i C_i \log(C_i/\bar{c}_i) + \sum_i \bar{c}_i \log(\bar{c}_i/d_i)$ gives the first identity in Equation (15). The gradient of $\mathrm{KL}(\boldsymbol{C}\|\mathrm{softmax}(\boldsymbol{z}^s))$ with respect to $\boldsymbol{z}^s$ is $\boldsymbol{d} - \boldsymbol{C}$, and taking expectations gives the second. For two agents, $c_p = \mathrm{sigm}(u) = \frac{1}{2} + \frac{1}{2}\tanh(u/2)$, so Equation (14) with $\alpha = 1$ yields $\mathbb{E}\,\mathrm{sigm}(u^*_{pq} + \eta_{pq}) - \frac{1}{2} = \kappa^{(1)}_{pq}\big(\mathrm{sigm}(u^*_{pq}) - \frac{1}{2}\big)$, which is Equation (16). The derivative of the loss with respect to the student gap is $\mathrm{sigm}(v) - \bar{c}_p$ in expectation and $\mathrm{sigm}(v) - c^*_p$ for the reference target; for v strictly between $\mathrm{logit}(\bar{c}_p)$ and u^*_{pq}, the two have opposite signs. □

Proof of Proposition 1. The relation is inverted if and only if $X_{pq} \geq \Delta_{pq}$. When $\Delta_{pq} \leq b_{pq}$, the bound equals one. Otherwise, with $t = \Delta_{pq} - b_{pq} > 0$ and any $\lambda > 0$, Markov's inequality gives $\mathbb{P}(X_{pq} - b_{pq} \geq t) \leq \exp(-\lambda t + \lambda^2\nu^2_{pq}/2)$, and $\lambda = t/\nu^2_{pq}$ gives the per-relation bound. Summing over relations and dividing by M gives Equation (17) by linearity of expectation. For strict orders, d_K is the fraction of discordant relations, which equals $(1-\tau)/2$ and satisfies the triangle inequality. □

Proof of the agent inequality in Section 4. With θ_{pq} the angle between $\boldsymbol{a}_p$ and $\boldsymbol{a}_q$, $\|\boldsymbol{a}_p - \boldsymbol{a}_q\|^2 = \|\boldsymbol{a}_p\|^2 + \|\boldsymbol{a}_q\|^2 - 2\|\boldsymbol{a}_p\|\|\boldsymbol{a}_q\|\cos\theta_{pq}$ and $\cos\theta_{pq} \leq \rho(\boldsymbol{A})$. If $\rho(\boldsymbol{A}) \geq 0$, the bound $2\|\boldsymbol{a}_p\|\|\boldsymbol{a}_q\| \leq \|\boldsymbol{a}_p\|^2 + \|\boldsymbol{a}_q\|^2$ gives $\|\boldsymbol{a}_p - \boldsymbol{a}_q\|^2 \geq (1-\rho(\boldsymbol{A}))(\|\boldsymbol{a}_p\|^2 + \|\boldsymbol{a}_q\|^2) \geq 2r^2(1-\rho(\boldsymbol{A}))$. If $\rho(\boldsymbol{A}) < 0$, the bound $\|\boldsymbol{a}_p\|\|\boldsymbol{a}_q\| \geq r^2$ gives $\|\boldsymbol{a}_p - \boldsymbol{a}_q\|^2 \geq 2r^2 - 2\rho(\boldsymbol{A})\,r^2$. The margin bound is the Cauchy–Schwarz inequality. □

Remark 1 (Value alignment can reverse an order). Let $N = 3$, $\boldsymbol{z}^* = (0, 0.2, 3)$, and perturb only agent 1 by $\delta_1 \sim \mathcal{N}(0,1)$. Then $\boldsymbol{c}^* \approx (0.045, 0.055, 0.900)$ and $\bar{\boldsymbol{c}} \approx (0.066, 0.054, 0.881)$: the target of value alignment ranks agent 1 above agent 2 although $z^*_1 < z^*_2$. For the relation $(1,2)$, $\eta_{12} = \delta_1$ is symmetric; it is inverted with probability 0.42, and with $\alpha = 2$ the expected update of RDC keeps its direction with $\kappa_{12} \approx 0.61$. In the extreme case $\boldsymbol{z}^* = (0, -0.1, 0)$ and $\boldsymbol{\delta} = (0, \pm 5, 0)$ with equal probability, $\boldsymbol{c}^* \approx (0.344, 0.312, 0.344)$ and $\bar{\boldsymbol{c}} \approx (0.253, 0.494, 0.253)$, while $\kappa_{12} \approx 1.8 \times 10^{-4}$: order alignment fades and keeps its direction. With many agents, independent $\delta_n \sim \mathcal{N}(0, s_n^2)$, and a normalizer $Z = \sum_k e^{z_k}$ that varies little, $\bar{c}_n = \mathbb{E}[e^{z_n}/Z] \approx e^{z^*_n + s_n^2/2}/\mathbb{E}Z$, so the target of value

alignment orders the agents by $z_n^* + s_n^2/2$. In a simulation with 64 agents, $z_1^* = 0$, $z_2^* = 0.5$, $s_1 = 1.5$, and $s_n = 0.3$ otherwise, the target of value alignment ranks agent 1 above agent 2, while $\kappa_{12} = 0.49$.

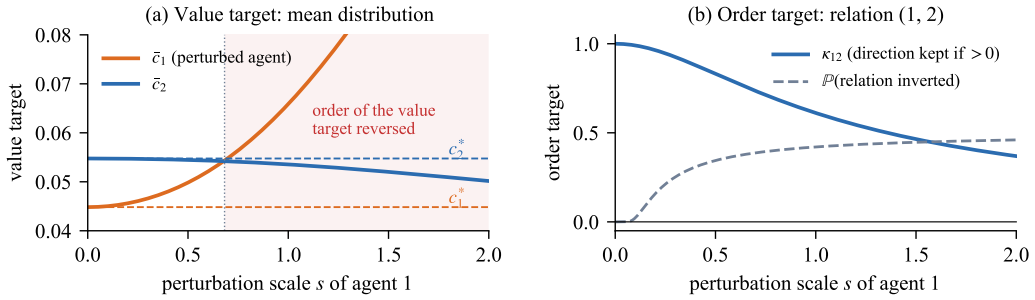


Figure 5: Value and order alignment under teacher noise, for three agents with reference affinities $\boldsymbol{z}^* = (0, 0.2, 3)$ and a perturbation $\delta_1 \sim \mathcal{N}(0, s^2)$ of agent 1 (Remark 1). **(a)** The target of value alignment, $\bar{\boldsymbol{c}}$, ranks agent 1 above agent 2 once $s > 0.68$, although $z_1^* < z_2^*$. **(b)** For order alignment, the inversion probability of relation $(1, 2)$ grows with s, while κ_{12} stays positive: the expected update keeps the reference direction and weakens.

Remark 2 (From per-relation decay to aggregate scaling). Suppose $b_{pq} = 0$ and $\nu_{pq} = \sigma c_{pq}$, where σ is a global noise scale and $c_{pq} > 0$ depends only on the relation; for isotropic noise injected into all embeddings, $c_{pq} \approx (\|\boldsymbol{a}_p - \boldsymbol{a}_q\|^2 + 2\|\boldsymbol{f}\|^2)^{1/2}$ to first order. Let $R_{pq} = \Delta_{pq}/c_{pq}$ be the normalized margin and $F(t)$ the fraction of relations with $R_{pq} \le t$, with $F(0) = 0$. Integrating by parts,

$$\frac{1}{M}\sum_{p<q} e^{-R_{pq}^2/(2\sigma^2)} = \int_0^\infty e^{-t^2/(2\sigma^2)}\, dF(t) = \int_0^\infty F(t)\,\frac{t}{\sigma^2}\, e^{-t^2/(2\sigma^2)}\, dt.$$

If $F(t) \le Ct^\beta$ for all $t \ge 0$, the substitution $x = t^2/(2\sigma^2)$ bounds the right-hand side by $C\, 2^{\beta/2}\, \Gamma(1+\beta/2)\, \sigma^\beta$. Each relation decays exponentially, while the average is at most of order σ^β, with β the small-margin exponent of F. If the normalized margins have a density h with $h(0) > 0$, the right-hand side equals $h(0)\sqrt{\pi/2}\,\sigma\,(1 + o(1))$ as $\sigma \to 0$, so $\beta = 1$.

Remark 3 (Local size of the value discrepancy). With $\boldsymbol{\delta} = \sigma\boldsymbol{\xi}$, $\boldsymbol{\xi}$ of fixed distribution with finite third moment and covariance $\boldsymbol{\Sigma}_\xi$, and $\boldsymbol{H}^* = \mathrm{diag}(\boldsymbol{c}^*) - \boldsymbol{c}^*\boldsymbol{c}^{*\top}$, a second-order expansion of the log-partition function gives, as $\sigma \to 0$,

$$\mathbb{E}\,\mathrm{KL}\big(\boldsymbol{c}^* \,\|\, \mathrm{softmax}(\boldsymbol{z}^* + \boldsymbol{\delta})\big) = \frac{\sigma^2}{2}\Big(\mathrm{tr}(\boldsymbol{H}^*\boldsymbol{\Sigma}_\xi) + \mathbb{E}[\boldsymbol{\xi}]^\top \boldsymbol{H}^* \mathbb{E}[\boldsymbol{\xi}]\Big) + O(\sigma^3).$$

The expected divergence between the reference target and a perturbed one therefore grows as σ^2, and the margins enter only through $\boldsymbol{H}^*$, which shrinks as $\boldsymbol{c}^*$ saturates. This is the $\sigma^{2.03}$ growth of the KL in Figure 8; the L1 distance between distributions is first order in σ and grows approximately linearly.

C Measurements for Section 4

Except for the training with teacher noise, all measurements use trained checkpoints and require no retraining.

Inverted relations (Table 7a). The bound of Proposition 1 is applied with the teacher affinities $\boldsymbol{z}^w$ as reference, zero offset, and the student deviation $\boldsymbol{\delta} = \boldsymbol{z}^s - \boldsymbol{z}^w$ on all pixels of a held-out batch ($N = 64$). The sub-Gaussian envelope $2\exp(-t^2/(2\hat{\nu}^2))$ with $\hat{\nu}^2 = 1.45$ dominates the measured tail of $\delta_p - \delta_q$ at $t = 2, 3, 4$ (exceedance $0.050, 0.008, 0.0015$ against $0.501, 0.089, 0.0079$). Relations are stratified by $\Delta_{pq}/\hat{\nu}$, with 394.8, 286.1, 266.6, and 76.8 million relations in the four strata.

Relation accuracy (Table 7b). On labeled images, the teacher predicts a pseudo-label for every pixel and agents are selected as in training. A relation (p, q) of pixel j is evaluated when the ground-truth labels satisfy $y_p = y_j \neq y_q$, and it is correct if $z_{jp}^w > z_{jq}^w$. We report the fraction of correct relations separately for pixels whose pseudo-label is correct and incorrect.

Training with teacher noise (Table 8). Gaussian noise is added to the teacher affinities z^w used by the inter-pixel term, with a standard deviation of 0.85 (low) and 1.70 (high), that is, one and two times the per-agent scale $\hat{\nu}/\sqrt{2}$ of the perturbation measured during training ($\hat{\nu}^2 = 1.45$, Table 16). The pseudo-labels of $\mathcal{L}_{reg}$ come from the unperturbed teacher, and KL and RDC share the agents, λ, and the schedule, so that the two runs differ only in the alignment.

Relation strength and coverage (Section 5.3). For every unlabeled pixel we compute the relation strength $S_j = \frac{1}{M}\sum_{p<q} |g_\alpha(u_{j,pq})|$ and compare pixels below and above the confidence threshold γ.

Perturbation seen by the loss. Table 16 summarizes the student–teacher perturbation $\boldsymbol{\delta} = z^s - z^w$ measured during training. The tail of $\delta_p - \delta_q$ lies well inside the sub-Gaussian envelope used in Table 7a. The components of $\boldsymbol{\delta}$ are correlated across agents through a component shared by all agents, which cancels in $\delta_p - \delta_q$: after removing it, the mean correlation drops from 0.237 to 0.0002. Proposition 1 requires no independence, and the mean offset of the perturbation enters its bound through the means b_{pq}.

Table 16: Perturbation $\boldsymbol{\delta} = z^s - z^w$ during training (ACDC 5%, $N = 64$, all pixels of a held-out batch).

Quantity	Value		
Sub-Gaussian proxy variance $\hat{\nu}^2$ of $\delta_p - \delta_q$		1.45	
$\mathbb{P}(\lvert\delta_p - \delta_q\rvert > t)$ at $t = 2, 3, 4$	0.050	0.008	0.0015
Envelope $2e^{-t^2/(2\hat{\nu}^2)}$ at $t = 2, 3, 4$	0.501	0.089	0.0079
Mean cross-agent correlation		0.237	
after removing the shared component		0.0002	
Normalized mean offset $\lvert\mathbb{E}\,\delta\rvert/\mathrm{sd}(\delta)$		0.865	

Table 17: Log–log slopes under injected noise (ACDC 5%, same agents).

Objective	Slope
KL (value)	1.998
Plackett–Luce KL (value)	1.966
Kendall (RDC, order)	0.885

Per-relation decay and aggregate scaling. Figure 6 injects Gaussian noise of scale σ into the teacher's pixel embeddings. The inversion probability of individual relations decays exponentially in Δ_{pq}^2/σ^2, the per-relation behavior of Proposition 1, while the average inversion rate grows as $\sigma^{0.89}$, the aggregate power law of Remark 2. Figure 7 shows the distribution of margins with orthogonal agents: the median margin is 1.41 and the 95th percentile 5.09, while the per-pixel minimum margin is small (median 0.063, Figure 9). Since Equation (17) averages over relations, the supervision of a pixel is governed by this whole distribution, and the few near-ties contribute little.

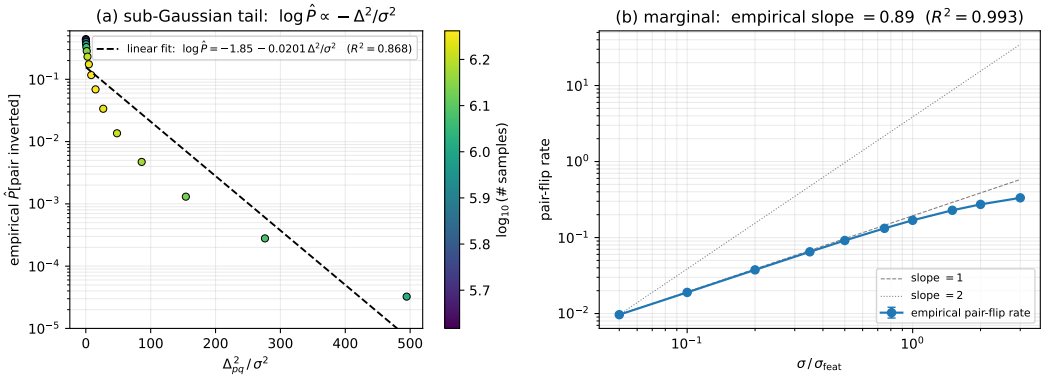


Figure 6: Inversion of relations under Gaussian noise injected into the teacher's pixel embeddings (ACDC 5%). **(a)** Empirical inversion probability of a relation against Δ_{pq}^2/σ^2 (semi-log scale); the linear fit (R^2=0.868) shows the exponential decay of individual relations. **(b)** Average inversion rate against the noise level (log–log scale); the slope of 0.89 (R^2=0.993) is the aggregate power law of Remark 2.

Loss under injected noise. Figure 8 records how the losses themselves change under the same injected noise. The KL grows as $\sigma^{2.03}$, the second-order behavior of Remark 3; the L1 distance grows approximately linearly; and the order loss grows as $\sigma^{0.89}$ and saturates, as it is bounded. Table 17

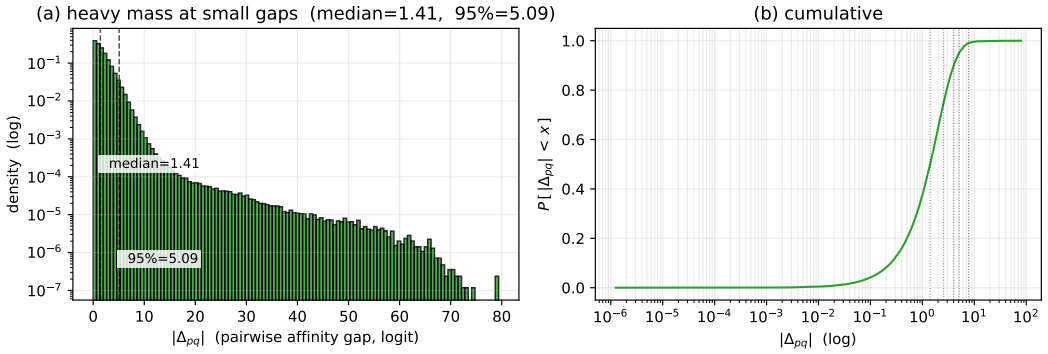


Figure 7: Distribution of the margins $|\Delta_{pq}|$ in logit space on ACDC (5%) with orthogonal agents (N=64). **(a)** Log-scale histogram (median 1.41, 95th percentile 5.09). **(b)** Cumulative distribution on log-scaled $|\Delta_{pq}|$.

repeats the comparison on identical agents: both value objectives, including the Plackett–Luce KL, grow approximately quadratically, and the order objective grows with slope 0.885.

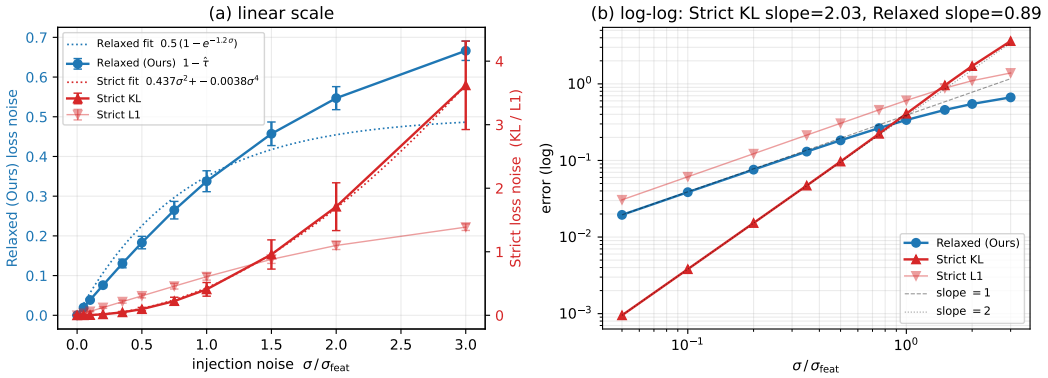


Figure 8: Deviation of the losses from their clean values under Gaussian noise injected into the teacher's pixel embeddings (ACDC 5%); σ is normalized by the feature standard deviation σ_{feat}. **(a)** Linear scale: order loss $1 - \hat{\tau}$ (left axis) and value losses KL and L1 (right axis). **(b)** Log–log scale.

D ALGORITHMS

Algorithm 1 is the orthogonal agent selection of RankMatch (Mai et al., 2024). It starts from a random pixel and repeatedly adds the pixel whose largest cosine similarity to the selected agents is the smallest. The indices are computed once per image from the teacher embedding and shared by teacher and student. Its cost is $O(HW \cdot N^2 \cdot C)$ per image.

Algorithm 1 Greedy orthogonal agent selection.

Require: Pixel embedding $\boldsymbol{F} \in \mathbb{R}^{C \times HW}$, number of agents N
Ensure: Agent matrix $\boldsymbol{A} \in \mathbb{R}^{C \times N}$

$\boldsymbol{A} \leftarrow \texttt{random_select}(\boldsymbol{F})$ ▷ one random pixel embedding
for $i = 1$ **to** $N - 1$ **do**
  $\mathcal{F}^- \leftarrow \{\boldsymbol{f} \in \boldsymbol{F} : \boldsymbol{f} \notin \boldsymbol{A}\}$ ▷ remaining pixels
  $\boldsymbol{S} \leftarrow \texttt{cosine_similarity}(\mathcal{F}^-, \boldsymbol{A})$ ▷ $\boldsymbol{S} \in \mathbb{R}^{(HW-i) \times i}$
  $\boldsymbol{s} \leftarrow \max(\boldsymbol{S}, \text{dim} = 1)$ ▷ largest similarity to the current agents
  $\boldsymbol{a} \leftarrow \mathcal{F}^-[\arg\min(\boldsymbol{s})]$
  $\boldsymbol{A} \leftarrow \boldsymbol{A} \cup \{\boldsymbol{a}\}$
end for

Algorithm 2 summarizes one training iteration.

Algorithm 2 One training iteration of RDC.

Require: Labeled batch $\{(\boldsymbol{x}_i^l, \boldsymbol{y}_i^l)\}$, unlabeled batch $\{\boldsymbol{x}_i^u\}$; teacher f_T ($\boldsymbol{\theta}_T$), student f_S ($\boldsymbol{\theta}_S$); augmentations aug, $\mathcal{A}ug$; weight λ, EMA momentum m
Ensure: Updated $\boldsymbol{\theta}_S$ and $\boldsymbol{\theta}_T$

Compute $\mathcal{L}_{sup}$ on the labeled batch (Equation (1))
Generate pseudo-labels $\hat{\boldsymbol{y}}_{ij}^u$ with the teacher (Equation (2)) and compute $\mathcal{L}_{reg}$ (Equation (3))
Select agent indices $\mathcal{I}$ from the teacher embedding $\boldsymbol{F}^w$ (Algorithm 1)
Gather $\boldsymbol{A}^w = \boldsymbol{F}^w[:, \mathcal{I}]$ and $\boldsymbol{A}^s = \boldsymbol{F}^s[:, \mathcal{I}]$
Compute $\boldsymbol{z}_{ij}^w = (\boldsymbol{A}^w)^\top \boldsymbol{f}_{ij}^w$ and $\boldsymbol{z}_{ij}^s = (\boldsymbol{A}^s)^\top \boldsymbol{f}_{ij}^s$
Compute $\mathcal{L}_{rdc}$ (Equations (7) and (8)) ▷ teacher affinities detached
Update $\boldsymbol{\theta}_S$ with $\mathcal{L} = \mathcal{L}_{sup} + \mathcal{L}_{reg} + \lambda \mathcal{L}_{rdc}$
$\boldsymbol{\theta}_T \leftarrow m\,\boldsymbol{\theta}_T + (1 - m)\,\boldsymbol{\theta}_S$

E Additional Analyses

E.1 Agent Selection

Table 18 compares four ways of choosing agents on real features: all pixels, N random pixels, the top N pixels by cumulative self-correlation, and orthogonal agent selection. Orthogonal agent selection yields both the largest per-pixel minimum margin $\Delta_{\min}$ and the highest Dice, whereas using all pixels collapses $\Delta_{\min}$ to zero, as the agent inequality of Section 4 predicts for near-duplicate agents; Figure 9 shows the full distributions. The pixel embeddings are far from isotropic: the ratio between the largest and smallest eigenvalues of their covariance is 235.5, and the normalized participation ratio is 0.270, against 1.0 for isotropic features. The advantage of orthogonal agent selection holds under this anisotropy. Figure 10 shows that different agents activate distinct regions of the image. Table 19 repeats the selection over five seeds: the selected agents return to the same locations, their activation regions overlap eight times more than those of random agents, and replacing them with random pixels costs 1.63 Dice and increases the run-to-run standard deviation from 0.15 to 0.37.

Table 18: Agent selection on ACDC (5%): minimum margin and Dice.

Agent selection	$\Delta_{\min}$	Dice
All pixels	0	88.17
Random	0.0557	88.82
Top-N	0.0122	89.13
Orthogonal	**0.0633**	**90.45**

Table 19: Stability of agent selection over five seeds (ACDC 5%, $N = 64$).

Measure	Orthogonal	Random
Activation-region IoU (32^2 grid)	0.267	0.033
Median nearest-agent displacement (px)	4.0	15.7
Std of agents per foreground class	0.9	–
Dice (mean $\pm$ std)	90.45 ± 0.15	88.82 ± 0.37

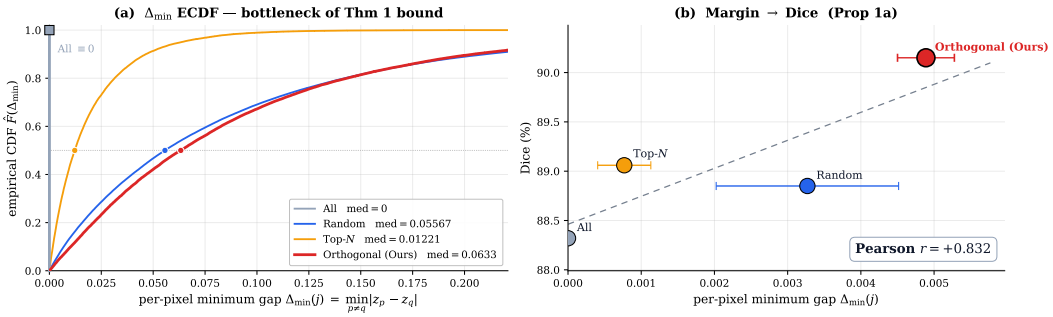


Figure 9: Margins and Dice across agent selection strategies on ACDC (5%). **(a)** Empirical distribution of the per-pixel minimum margin $\Delta_{\min}(j) = \min_{p \neq q} |z_p - z_q|$. **(b)** Mean $\Delta_{\min}$ against Dice (Pearson $r = 0.819$).

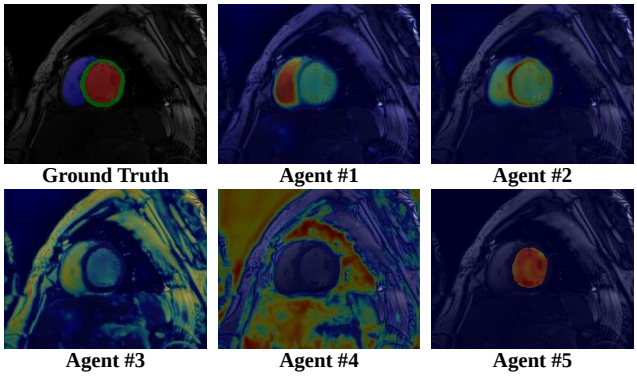


Figure 10: Activation maps of agents obtained by orthogonal agent selection.

E.2 Per-class Results on Synapse

The Synapse dataset is highly imbalanced: the liver covers 71.23% of the foreground pixels, while the four smallest organs (aorta, gallbladder, left and right kidneys) together cover less than 3.5%. Table 20 reports per-class Dice with one labeled case, the most extreme setting in our experiments. RDC outperforms all baselines on every organ, and its gains concentrate on the organs where ABD is weakest: pancreas ($8.6 \rightarrow 13.5$, $+57\%$ relative), stomach ($21.2 \rightarrow 28.5$, $+34\%$), left kidney ($14.8 \rightarrow 19.5$, $+32\%$), and gallbladder ($11.7 \rightarrow 15.1$, $+29\%$), against $+3.5\%$ to $+5.5\%$ on the four organs ABD already segments well. These difficult organs are where the teacher is least confident: $\mathcal{L}_{rdc}$ is applied to all pixels regardless of confidence, so their pixels still receive inter-pixel supervision when $\mathcal{L}_{reg}$ discards them through the threshold γ.

Table 20: Per-class Dice (%) on **Synapse** with one labeled case. The second row gives the foreground-pixel ratio of each class. RDC outperforms all baselines on every class, with the largest relative gains on the organs where ABD is weakest (pancreas, stomach, left kidney, gallbladder).

Method *ratio (%)*	Liver *71.23*	Stomach *15.20*	Pancreas *6.20*	Spleen *3.91*	Aorta *2.61*	Gallbladder *0.56*	Kid (R) *0.18*	Kid (L) *0.12*	Mean —
Sup.-only	33.8	6.9	1.5	8.7	14.7	9.1	5.6	5.3	10.7
DTC (Luo et al., 2021a)	50.4	8.2	3.3	10.5	20.6	10.3	10.5	6.2	15.0
ABD (Chi et al., 2024)	76.0	21.2	8.6	64.1	62.5	11.7	69.9	14.8	41.1
RDC (Ours)	**79.2**	**28.5**	**13.5**	**67.6**	**64.7**	**15.1**	**72.9**	**19.5**	**45.1**

E.3 Behaviour across Regimes

Table 21 collects the settings in which order alignment helps less. Each follows from the analysis: order alignment transfers the teacher's relations, so it needs an informative teacher; it carries no class semantics, so it complements $\mathcal{L}_{reg}$ rather than replacing it; and the number of agents trades a finer description of each pixel against smaller margins.

Table 21: Regimes in which order alignment helps less (Dice %).

Regime	Observation	Interpretation
Teacher close to random (ACDC, one labeled case)	83.91	needs an informative teacher order
Order alignment without $\mathcal{L}_{reg}$ (Table 6)	77.91 / 79.03	carries no class semantics
N far from 64 (Table 12)	-0.90 to -1.80	finer description vs. smaller margins
Large weight $\lambda = 0.5$ (Table 12)	-1.14 to -1.64	weaker complementarity with $\mathcal{L}_{reg}$

E.4 Coverage of Pixel-level and Inter-pixel Supervision

$\mathcal{L}_{reg}$ is restricted to pixels whose teacher confidence exceeds γ, whereas $\mathcal{L}_{rdc}$ is applied to every pixel (Section 3.4). Figure 11 visualizes this difference on three unlabeled ACDC slices. Teacher confidence is high over most of the cardiac region and drops in narrow bands along organ boundaries. These bands are excluded from $\mathcal{L}_{reg}$ and supervised only by $\mathcal{L}_{rdc}$; they trace the myocardium contour and the interfaces between ventricles and cover 4–8% of the cropped region. Boundary pixels are both the hardest to classify and the most influential for Dice, since a small displacement of a boundary changes the overlap more than the same displacement in the interior.

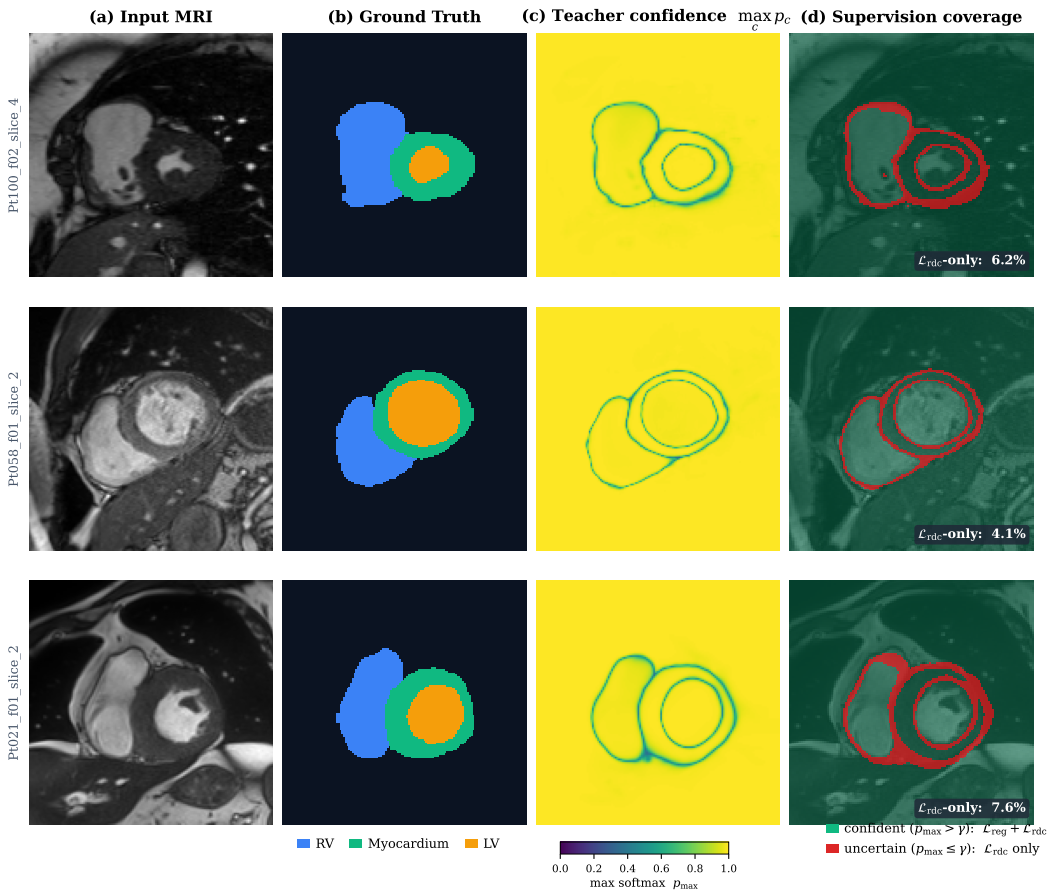


Figure 11: Supervision coverage on three unlabeled ACDC slices (5% labeled). **(a)** Input MRI, cropped to the heart. **(b)** Ground truth, shown for reference only: blue = RV, green = myocardium, orange = LV. **(c)** Teacher confidence $\max_c p_c$: confidence drops below γ=0.95 in narrow boundary bands (cyan). **(d)** Coverage: green pixels ($p_{\max} > \gamma$) receive both $\mathcal{L}_{reg}$ and $\mathcal{L}_{rdc}$; red pixels ($p_{\max} \leq \gamma$) are excluded from $\mathcal{L}_{reg}$ and supervised only by $\mathcal{L}_{rdc}$. The red bands trace organ boundaries and cover 4–8% of the cropped image.

F Extended Related Work

F.1 Consistency Regularization in SSMIS

The teacher–student framework, popularized by Mean Teacher (Tarvainen & Valpola, 2017) and FixMatch (Sohn et al., 2020), has been adapted to medical image segmentation along three lines. The first improves the reliability of the pseudo-labels: UA-MT (Yu et al., 2019) uses Monte-Carlo-dropout uncertainty, SASSNet (Li et al., 2020) adds shape awareness through signed distance maps, ASE-Net (Lei et al., 2022) combines adversarial consistency with dynamic convolution, and CauSSL (Miao et al., 2023) disentangles content and style with structural causal models. The second uses multiple views or models: MC-Net+ (Wu et al., 2021) trains several decoders with mutual consistency, MCF (Wang et al., 2023) lets two models correct each other, and CPS (Chen et al., 2021) uses each network's predictions as pseudo-labels for the other. The third designs augmentations and perturbations: PS-MT (Liu et al., 2022) combines perturbed and strict mean teachers, BCP (Bai et al., 2023) copies and pastes between labeled and unlabeled images, ABD (Chi et al., 2024) displaces features bidirectionally, GapMatch (Huang et al., 2025) unifies data- and model-level perturbations, UniMatch (Yang et al., 2022) revisits weak-to-strong consistency, and VQ-Seg (Yang et al., 2026) perturbs vector-quantized tokens. All of these methods supervise the student with pixel-level pseudo-labels, which RDC complements with inter-pixel supervision.

F.2 Inter-pixel Supervision for Dense Prediction

Pixel-level contrastive learning pulls positive pairs together and pushes negative pairs apart, across images in fully supervised segmentation (Wang et al., 2021) and with a class-wise memory bank in the semi-supervised setting (Alonso et al., 2021). For medical images, SS-Net (Wu et al., 2022) imposes smoothness and class separation, GraphCL (Wang et al., 2024) applies contrastive losses on graph clusters, momentum contrastive voxel distillation (You et al., 2022) extends contrastive learning to 3D, cross-level contrastive learning (Zhao et al., 2022) adds constraints across feature levels, and ARCO (You et al., 2023) reduces the variance of the contrastive estimator. In the semi-supervised setting, the positive and negative pairs are built from the model's own predictions. RankMatch (Mai et al., 2024) describes each pixel by its affinities to agents obtained by orthogonal agent selection and enforces consistency on these affinities between teacher and student. RDC shares this pixel–agent representation and changes the alignment from the values of the affinities to their order; Table 22 compares the two methods.

F.3 SAM Adaptations for Medical Image Segmentation

The Segment Anything Model (SAM) (Kirillov et al., 2023) has been adapted to medical images in two ways. The first fine-tunes SAM for the medical domain: MedSAM (Ma et al., 2024) fine-tunes it on over one million medical image–mask pairs, SAMed (Zhang & Liu, 2023) applies low-rank adaptation, adapter-based variants (Wu et al., 2025) target specific modalities, and SAM-Med3D (Wang et al., 2025) extends SAM to volumetric segmentation. The second integrates SAM into semi-supervised pipelines: SemiSAM (Zhang et al., 2024) uses SAM predictions as additional pseudo-labels, SemiSAM+ (Zhang et al., 2025) improves the prompts, UP-SAM (Lu et al., 2024) adds uncertainty-aware adaptation, Li et al. (2025) stitch, fine-tune, and retrain SAM for 3D segmentation, and ESDE (Zhao et al., 2026) distills a full SAM teacher into an efficient SAM student through a KL divergence. In these pipelines the foundation model provides a stronger teacher or backbone; RDC changes the alignment and therefore applies to them as a plug-in (Section A.7).

F.4 Ranking Objectives

Kendall's τ (Kendall, 1938) measures the agreement between two rankings as the proportion of concordant minus discordant pairs, and Spearman's ρ (Spearman, 1961) compares rank differences. To train ranking objectives by gradient descent, RankNet (Burges et al., 2005) introduces a pairwise logistic loss, LambdaRank (Burges et al., 2006) weights gradients by ranking metrics, ListNet (Cao et al., 2007) and ListMLE (Xia et al., 2008) take listwise approaches, and NeuralSort (Grover et al., 2019), SoftSort (Prillo & Eisenschlos, 2020), and fast differentiable sorting (Blondel et al., 2020) relax the sorting operator. Ranking losses are also used in metric learning (Schroff et al., 2015)

Table 22: RankMatch and RDC. Both methods build on the standard teacher–student framework $\mathcal{L}_{sup} + \mathcal{L}_{reg} + \lambda\mathcal{L}_{aux}$.

Level	Dimension	RankMatch (Mai et al., 2024)	RDC (this work)
Strategy	Focus	richer inter-pixel supervision	reliable inter-pixel supervision
	Alignment	value alignment	order alignment
Instantiation	Objective	formalizable as a Plackett–Luce KL	differentiable Kendall's τ
	Reference pixels (agents)	orthogonal agent selection	adopted from RankMatch
Analysis	Theory	–	Lemma 1, Theorem 1, Proposition 1
	Setting	natural images (VOC, Cityscapes)	medical images (ACDC, LA, Pancreas-NIH, Synapse)
Result	Dice on the same agents (ACDC-5% / LA-5%)	88.71 / 89.39	90.45 / 90.01

and depth ordering (Rezaeirowshan et al., 2016). DiffKendall (Zheng et al., 2023) approximates Kendall's τ by replacing the sign function with a hyperbolic tangent, and the same surrogate improves logit distillation from a trained teacher (Guan et al., 2024). RDC uses this surrogate on pixel–agent affinities, where the teacher is trained jointly with the student, and Section 4 analyzes how the teacher's perturbations reach the student's update. Table 5 compares Kendall's τ with RankNet, a pairwise hinge loss, Spearman's ρ, and SoftSort on the same agents.

F.5 Learning under Label Noise

Noise-tolerant losses are a classical response to label noise. Symmetric losses, whose values over all labels sum to a constant, make risk minimization tolerant to symmetric label noise (Ghosh et al., 2017); mean absolute error is symmetric but converges slowly, generalized cross-entropy (Zhang & Sabuncu, 2018) interpolates between cross-entropy and mean absolute error, normalized losses (Ma et al., 2020) are tolerant by construction, and symmetric cross-entropy (Wang et al., 2019) adds a reverse cross-entropy term. The student term of RDC is the sigmoid loss, which is symmetric (Section 3.3), applied to pairwise relational pseudo-labels. Semi-supervised learning also controls the quality of pseudo-labels through confidence, with fixed (Sohn et al., 2020), curriculum (Zhang et al., 2021), self-adaptive (Wang et al., 2022), or soft (Chen et al., 2023) thresholds; these rules discard or down-weight uncertain pixels, on which RDC still provides inter-pixel supervision. Loss correction with a transition matrix (Patrini et al., 2017) requires an estimate of the noise structure. Theoretical analyses of self-training (Wei et al., 2020) show that consistency improves sample efficiency under expansion assumptions; Section 4 studies how the form of the alignment shapes the update that an imperfect teacher induces.