Tell2Adapt: A Unified Framework for Source Free Unsupervised Domain Adaptation via Vision Foundation Model

Appendix

This appendix complements the main paper by providing extended quantitative evaluations and detailed implementation specifications to ensure full reproducibility. While the main paper focuses primarily on DICE-based comparisons in abdominal targets, this appendix further reports ASD and elaborates on the prompt regularization mechanism that supports the Tell2Adapt framework. The appendix is organized as follows:

- **Extended Quantitative Evaluation (Section A):** We present the complete ASD for all abdominal organs in both the MR→CT and CT→MR directions, together with ASD from the ablation study. These results reinforce the effectiveness of Tell2Adapt in achieving accurate and anatomically consistent segmentations.
- **Details in Ablation Study for CAPR (Section B):** We describe the procedure for constructing chaos prompts used in the ablation study and provide the full list of chaos prompts. This section details how perturbations are introduced to assess the robustness of CAPR.
- **Reproducibility Details (Section C):** To enable exact replication, we include the meta-prompt used in CAPR for LLM and the complete set of input prompts for all anatomical targets.

A. Extended ASD Evaluation on Abdominal Targets

To complement DICE reported in the main paper, this section provides a comprehensive analysis of ASD performance for all abdominal targets. We report full ASD for both MR→CT and CT→MR directions, following the same evaluation protocol used in the main paper. These results offer a finer-grained assessment of boundary accuracy, which is particularly important for small and geometrically complex organs where DICE alone may not fully capture segmentation quality.

In addition to the adaptation results, we also include ASD from the ablation study to further elucidate the contribution of CAPR and VPR. Together, these quantitative evaluations provide a comprehensive assessment of Tell2Adapt's effectiveness in achieving anatomically precise segmentations under severe cross-modality domain shifts.

A.1. Adaptation Results in ASD for Abdominal Targets

Due to space limitations, Table 2 in the main paper reports only DICE for abdominal segmentation. For completeness, we provide the corresponding ASD in Table 7. All ASD are computed over the same test subjects used in the main paper, following identical evaluation protocols and spacing normalization.

As shown in Table 7, Tell2Adapt achieves consistently strong boundary accuracy across all abdominal organs. In MR→CT, our framework attains a mean ASD of 1.5 mm, representing a substantial reduction compared to Baseline and nearly matching Supervised of 1.4 mm. This strong performance is further reflected in the per-organ results, where struc-

Table 7. ASD (mm, mean ± std) of segmentation results on abdominal targets.

MR→CT																	
Methods	**VFM**	**Spleen**	**R. K**	**L. K**	**Gallb**	**Esoph**	**Liver**	**Stom**	**Aortic**	**V-Cava**	**Pancr**	**R. Adre**	**L. Adre**	**Duod**	**Bladder**	**Pros/Uter**	**mASD**
Baseline	×	17.5±19.2	17.2±18.5	15.5±17.8	23.1±30.7	7.8±9.6	11.4±7.2	15.3±14.6	8.9±9.9	6.9±11.2	9.1±10.7	4.7±3.4	4.8±4.2	10.3±9.7	N/A	N/A	13.0±4.8
Supervised	×	1.1±0.3	1.1±0.8	1.2±1.4	1.2±2.4	0.8±0.7	2.3±0.5	1.9±1.6	1.1±1.3	1.0±0.4	1.4±1.0	0.7±0.7	0.7±0.7	1.6±1.1	1.9±1.9	2.7±7.7	1.4±0.6
UPL [21]	×	**79.0±19.9**	87.2±15.2	72.4±17.6	N/A	14.8±17.6	37.0±20.6	31.6±21.5	23.7±12.8	50.2±15.4	29.3±16.7	N/A	N/A	54.4±23.4	N/A	N/A	48.0±23.6
ProtoContra [24]	×	99.5±25.5	72.6±16.3	**67.9±18.4**	94.1±27.6	34.4±26.9	75.6±30.4	59.4±19.0	24.8±16.3	37.1±15.1	30.8±15.1	N/A	N/A	45.8±18.7	N/A	N/A	58.4±24.6
IAPC [3]	×	92.3±80.2	**65.2±19.9**	69.8±21.2	**82.0±37.9**	25.8±14.7	54.4±31.0	77.6±47.8	29.9±20.0	27.7±15.5	54.2±32.7	N/A	N/A	45.3±25.0	**31.1±18.2**	**31.2±19.7**	54.9±22.2
DFG [8]	✓	98.5±21.0	77.0±21.1	72.5±9.4	N/A	10.7±8.6	67.3±24.5	54.0±16.6	20.5±6.3	30.5±7.9	30.5±15.5	N/A	N/A	34.4±13.9	N/A	N/A	49.6±27.1
IPLC [26]	✓	N/A	N/A	N/A	N/A	N/A	**12.6±4.4**	40.9±13.5	50.9±16.3	52.8±11.3	39.6±11.3	N/A	N/A	N/A	N/A	N/A	39.4±14.4
SRPL [14]	✓	N/A	N/A	N/A	N/A	**2.0±3.1**	31.2±30.0	**15.3±18.2**	**5.6±7.4**	**20.5±7.8**	**10.1±11.4**	N/A	N/A	**29.1±24.8**	N/A	N/A	**16.3±10.4**
Ours	✓	**1.1±0.5**	**1.5±4.1**	**1.2±1.3**	**2.3±6.5**	**1.0±1.1**	**2.8±1.4**	**2.0±1.2**	**1.0±0.7**	**1.0±0.4**	**1.3±1.0**	**0.6±0.3**	**0.7±0.7**	**2.1±3.6**	**1.8±2.6**	**1.4±1.2**	**1.5±0.6**

CT→MR															
Methods	**VFM**	**Spleen**	**R. K**	**L. K**	**Gallb**	**Esoph**	**Liver**	**Stom**	**Aortic**	**V-Cava**	**Pancr**	**R. Adre**	**L. Adre**	**Duod**	**mASD**
Baseline	×	22.0±28.4	10.5±15.4	6.1±7.0	5.3±6.1	15.3±20.0	9.4±12.2	8.7±7.4	16.2±15.8	12.6±19.2	9.0±8.6	1.5±0.3	2.8±2.4	15.8±13.8	10.4±5.6
Supervised	×	1.0±0.1	1.3±1.4	0.9±0.4	1.3±1.3	1.9±1.6	2.2±0.9	1.8±0.8	1.3±1.1	0.9±0.4	1.5±0.5	1.1±0.6	1.4±1.0	3.1±1.7	1.5±0.6
UPL [21]	×	89.5±14.6	**71.9±18.7**	62.9±25.4	N/A	N/A	53.4±20.8	61.1±25.5	14.3±19.7	N/A	33.3±24.8	N/A	N/A	N/A	55.2±23.0
ProtoContra [24]	×	74.3±21.1	89.8±19.4	82.4±16.7	N/A	59.0±41.6	95.5±33.9	81.6±22.1	63.4±23.7	50.1±23.9	48.5±25.9	**56.8±19.5**	N/A	N/A	70.1±16.0
IAPC [3]	×	**71.5±48.1**	82.4±47.6	85.0±65.0	N/A	30.5±16.9	62.7±49.4	96.1±56.4	35.1±32.3	**31.8±20.4**	74.8±39.1	N/A	N/A	N/A	63.3±23.5
DFG [8]	✓	99.9±27.0	83.8±3.3	80.9±17.0	N/A	**18.0±17.3**	76.1±35.8	85.4±22.6	24.4±29.5	57.2±23.1	47.5±22.4	N/A	N/A	N/A	63.7±27.0
IPLC [26]	✓	N/A	N/A	N/A	N/A	65.6±14.5	**14.3±5.5**	31.9±12.0	42.8±11.1	47.0±7.4	25.7±5.5	N/A	N/A	41.4±6.3	38.4±15.2
SRPL [14]	✓	N/A	N/A	**4.4±13.3**	N/A	N/A	16.9±22.7	**8.1±9.3**	**5.3±12.0**	N/A	**10.5±11.5**	N/A	N/A	**10.5±11.2**	**9.3±4.1**
Ours	✓	**1.3±1.1**	**0.9±0.4**	**0.9±0.3**	**5.0±10.1**	**1.3±1.0**	**1.9±0.7**	**1.5±1.4**	**1.0±0.7**	**1.1±0.5**	**1.2±0.6**	**1.1±0.6**	**1.3±0.6**	**2.3±1.1**	**1.6±1.1**

Note: An ASD of N/A indicates a complete prediction failure for the corresponding targets.

turally diverse organs such as spleen (1.1 mm), esophagus (1.0 mm), and pancreas (1.3 mm) exhibit highly accurate boundary localization.

The robustness of Tell2Adapt is further demonstrated in CT→MR. Here, Tell2Adapt achieves a mean ASD of 1.6 mm, again closely aligned with Supervised of 1.5 mm. Notably, challenging small and elongated structures, including the right adrenal gland (1.1 mm) and aorta (1.0 mm) are segmented with high boundary precision.

Overall, the ASD results reinforce the findings derived from DICE. Tell2Adapt not only delivers accurate region-level segmentation but also preserves detailed boundary structure across all abdominal targets. This consistent performance across both MR→CT and CT→MR, on one of the most complex multi-organ abdominal benchmarks, highlights the strong generalization capability of Tell2Adapt and its effectiveness in overcoming severe domain shifts.

A.2. Ablation Results in ASD for Abdominal Targets

As stated in Section 4.2, we provide the corresponding ASD for our full ablation study on the abdominal targets. This complements DICE results presented in Table 6, and the results detailed in Table 8 mirror and reinforce the trends observed for DICE.

Consistent with the DICE findings, CAPR emerges as the most influential method in improving boundary accuracy. Without CAPR, chaos prompts lead to a marked deterioration in boundary accuracy, with mean ASD increasing from 1.8 mm to 2.8 mm in MR→CT and from 1.9 mm to 3.5 mm in CT→MR. As illustrated by the comparison between Ours w/o CAPR (Chaotic Prompts) and Ours (Chaotic Prompts), several organs exhibit substantially larger surface distances, indicating pronounced boundary misalignment. This demonstrates CAPR's ability to stabilize VFM guidance by reconstructing coherent and semantically aligned prompts, even when the original prompts are noisy and ambiguous.

Beyond its corrective role, CAPR also provides measurable optimization benefits. When comparing Ours (Normal Prompts) to Ours w/o CAPR (Normal Prompts), the mean ASD decreases notably from 2.4 mm to 1.5 mm in MR→CT and from 2.8 mm to 1.6 mm in CT→MR, demonstrating that CAPR improves boundary precision even when the initial prompts are already well-formed. This further confirms that CAPR not only rescues performance under chaos prompts but also improves semantic grounding under normal conditions. VPR provides an additional refinement by removing anatomically implausible components that can yield high ASD even when DICE appears reasonable. The increase in ASD from 1.5 mm to 2.8 mm in MR→CT and from 1.6 mm to 3.2 mm in CT→MR, when comparing Ours w/o VPR (Normal Prompts) to Ours (Normal Prompts) underscores the importance of incorporating anatomical priors. By eliminating false positives and enforcing shape plausibility, VPR enhances the boundary accuracy of Tell2Adapt.

Overall, the ASD ablation results validate the complementary effects of CAPR and VPR. CAPR provides robustness against prompt variability and enhances semantic alignment, while VPR enforces anatomical plausibility and boundary correctness. Their combined impact explains the substantial ASD reductions achieved by Tell2Adapt across both MR→CT and CT→MR, paralleling the improvements previously observed in DICE.

Table 8. ASD (mm, mean ± std) of segmentation results in the ablation study on abdominal targets.

MR→CT																
Methods	**Spleen**	**R. K**	**L. K**	**Gallb**	**Esoph**	**Liver**	**Stom**	**Aortic**	**V-Cava**	**Pancr**	**R. Adre**	**L. Adre**	**Duod**	**Bladder**	**Pros/Uter**	**mASD**
Baseline	17.5±19.2	17.2±18.5	15.5±17.8	23.1±30.7	7.8±9.6	11.4±7.2	15.3±14.6	8.9±9.9	6.9±11.2	9.1±10.7	N/A	N/A	10.3±9.7	N/A	N/A	13.0±4.8
Ours w/o CAPR (Chaotic Prompts)	1.7±1.1	N/A	2.5±0.5	2.6±1.4	2.6±0.8	N/A	N/A	2.4±0.7	5.4±9.3	N/A	1.7±0.4	N/A	3.9±2.4	N/A	2.6±2.5	2.8±1.1
Ours w/o CAPR (Normal Prompts)	1.5±0.9	**2.2±0.4**	2.1±1.5	**2.6±0.8**	1.7±0.9	3.3±1.5	2.7±0.4	2.0±1.9	4.7±2.6	1.9±1.1	1.5±0.4	2.1±1.4	**2.2±1.9**	2.9±1.1	2.4±0.7	2.4±8
Ours w/o VPR (Normal Prompts)	2.1±1.5	2.4±1.3	2.4±2.1	2.9±1.8	2.5±0.7	3.5±3.0	3.7±2.9	2.2±1.7	5.3±4.8	2.6±1.8	1.5±1.2	2.4±0.9	3.1±1.7	3.4±2.7	2.6±3.3	2.8±0.9
Ours (Chaotic Prompts)	**1.2±0.4**	2.3±0.1	**1.3±0.4**	2.8±2.4	**1.5±0.5**	**2.3±0.2**	**2.1±5.5**	**1.4±0.9**	**1.5±0.2**	**0.9±0.4**	**0.9±4.2**	**1.0±4.4**	3.8±0.9	**1.7±2.0**	**1.8±1.5**	**1.8±0.8**
Ours (Normal Prompts)	**1.1±0.5**	**1.5±4.1**	**1.2±1.3**	**2.3±6.5**	**1.0±1.1**	**2.8±1.4**	**2.0±1.2**	**1.0±0.7**	**1.0±0.4**	**1.3±1.0**	**0.6±0.3**	**0.7±0.7**	**2.1±3.6**	**1.8±2.6**	**1.4±1.2**	**1.5±0.6**

CT→MR														
Methods	**Spleen**	**R. K**	**L. K**	**Gallb**	**Esoph**	**Liver**	**Stom**	**Aortic**	**V-Cava**	**Pancr**	**R. Adre**	**L. Adre**	**Duod**	**mASD**
Baseline	22.0±28.4	10.5±15.4	6.1±7.0	5.3±6.1	15.3±20.0	9.4±12.2	8.7±7.4	16.2±15.8	12.6±19.2	9.0±8.6	1.5±0.3	2.8±2.4	15.8±13.8	10.4±5.6
Ours w/o CAPR (Chaotic Prompts)	2.7±1.1	N/A	2.5±1.3	5.6±0.7	2.7±0.3	N/A	N/A	2.4±0.7	5.4±9.3	N/A	2.7±0.9	N/A	3.8±1.4	3.5±1.2
Ours w/o CAPR (Normal Prompts)	1.9±0.8	2.0±1.6	1.7±1.4	5.3±3.1	**2.2±1.4**	2.5±0.4	2.7±2.1	2.0±1.8	4.7±3.5	3.1±2.4	2.5±1.4	2.4±1.1	**2.9±1.5**	2.8±1.0
Ours w/o VPR (Normal Prompts)	2.4±0.7	2.5±1.9	2.2±2.1	5.3±4.4	2.6±1.2	3.1±2.7	2.9±1.4	2.3±2.1	5.3±4.1	3.7±2.8	2.9±1.4	2.8±0.5	3.4±2.9	3.2±1.0
Ours (Chaotic Prompts)	**1.4±0.2**	**1.5±0.1**	**0.9±0.4**	**4.6±1.8**	2.7±0.1	**1.0±0.0**	**1.7±0.4**	**1.4±0.5**	**1.5±0.8**	**1.6±0.4**	**1.7±0.4**	**1.2±2.3**	3.8±1.4	**1.9±1.1**
Ours (Normal Prompts)	**1.3±1.1**	**0.9±0.4**	**0.9±0.3**	**5.0±10.1**	**1.3±1.0**	**1.9±0.7**	**1.5±1.4**	**1.0±0.7**	**1.1±0.5**	**1.2±0.6**	**1.1±0.6**	**1.3±0.6**	**2.3±1.1**	**1.6±1.1**

Note: An ASD of N/A indicates a complete prediction failure for the corresponding targets.

B. Details on Chaos Prompt in Ablation Study for CAPR

To strictly evaluate the robustness of CAPR against noisy real-world prompts, we introduced a systematic chaos prompt generation method. Unlike standard prompts, these prompts are intentionally corrupted with varying degrees of typographical errors and syntactic scrambling to simulate user input instability. This section details the mathematical formulation of these perturbations and provides the exact set of chaos prompts used in our ablation study to ensure the reproducibility of our robustness benchmarks.

B.1. Chaos Prompts Generation Methodology

To systematically evaluate the robustness of our method against noisy and corrupted text prompts, we introduce a controlled method for perturbing text prompts. This method generates prompt variations with a quantifiable chaos score to simulate real-world scenarios where prompts may contain typos, grammatical errors, or incomplete information. We inject noise through three primary categories of perturbations, with error rates dynamically scaled proportional to the target chaos level:

- **Typographical Errors:** These simulate common spelling mistakes via character-level operations, including substitution (replacing a character with a random letter), transposition (swapping adjacent characters), and insertion (adding a random character), governed by the rate r_{spell}.
- **Syntactic Disruption:** We employ word order shuffling, where the positions of words within a prompt are randomly swapped to disrupt the grammatical structure, governed by the rate r_{shuffle}.
- **Information Loss:** We explicitly introduce character deletion, where characters are randomly removed from the text to simulate incomplete input, governed by the rate r_{remove}.

We quantify the chaos level using the normalized Levenshtein distance between the original and perturbed prompts:

$$S(P_{\text{pert}}) = \min\left(100, \frac{L(P_{\text{orig}}, P_{\text{pert}})}{\max(|P_{\text{orig}}|, |P_{\text{pert}}|)} \times 100\right) \tag{5}$$

where $L(\cdot,\cdot)$ denotes the Levenshtein distance, $S(P_{\text{pert}})$ denotes the chaos score, P_{orig} and P_{pert} represent the original and perturbed prompts respectively, and $|\cdot|$ denotes the string length. The chaos score ranges from 0 (identical to original) to 100 (maximum perturbation). For each target chaos level $\tau \in \{5, 15, 30, 50, 75\}$, we employ an iterative optimization strategy. For each level, we generate 50 candidate versions and select the one whose actual chaos score is closest to the target:

$$P_\tau^* = \arg\min_{P \in \mathcal{C}_\tau} |S(P_{\text{pert}}) - \tau| \tag{6}$$

where $\mathcal{C}_\tau$ is the set of candidate prompts generated for the target level τ, and $S(P_{\text{pert}})$ computes the chaos score of prompt P. The perturbation rates for each operation are dynamically adjusted based on the target chaos level:

$$r_{\text{spell}} = 0.5 \times \frac{\tau}{100} \tag{7}$$

$$r_{\text{shuffle}} = 0.7 \times \frac{\tau}{100} \tag{8}$$

$$r_{\text{remove}} = 0.2 \times \frac{\tau}{100} \tag{9}$$

This method allows us to systematically evaluate model performance degradation across different levels of prompt corruption, ranging from minor typos (Chaos 5-15) to severely corrupted prompts (Chaos 50-75). For the ablation study presented in the main paper, we set the target chaos level to $\tau = 75$ across all generated prompts, and the source code to generate chaos prompts is provided in the supplementary materials. The specific chaos prompts used in our experiments are explicitly listed in Table 9.

B.2. Chaos Prompt in Ablation Study

In this section, we provide a comprehensive list of perturbed text prompts used in the ablation study for CAPR. These chaos prompts were constructed to introduce a diverse range of textual noise to validate the necessity of CAPR. By documenting the exact strings in Table 9, researchers can replicate the specific noise conditions under which Tell2Adapt was evaluated. The table lists the chaos prompts used in both the MR→CT and CT→MR directions for all abdominal targets.

C. Meta-Prompt in CAPR and Full Prompts for All Targets

The Tell2Adapt's performance is guided by the quality of the prompts provided to the VFM. To ensure full transparency and allow for exact replication of our results, this section provides the meta-prompt in CAPR and the complete list of input prompts used for all anatomical targets evaluated in Section 4. The prompts are organized by targets and domain adaptations in Table 10, Table 11, and Table 12.

Table 9. Chaos prompt for ablation study in Tell2Adapt.

Target	**Chaos Prompt in MR→CT**	**Chaos Prompt in CT→MR**
Spleen	abdominal in CT Spleen.	pSleen MR in abdomianl.
Right Kidney	Right idney in abdominal CT.	MR kidey in Right abdominal.
Left Kidney	Left abdominl kdne in CT.	Lewt kdiney in abdominl MR.
Gallbladder	Gallbladedr in abdominal CT.	Gallbladder in abdomital MR.
Esophagus	in CT abdominal Esophagus.	abdominal in MR Esophagus.
Liver	Liiver in abdominal CT.	abdminal in MR Liver.
Stomach	abdominal in CT Stomach.	in MR xbdominal Stomach.
Aorta	in CT abdominal Aorta.	MR in Aorta abdominal.
Vena-Cava	cava in abdominal CT Vedna.	Vena cav in abdominal MR.
Pancreas	Pancreas CT. in bdominal.	Pancroeas in abdominal MR.
Right Adrenal gland	glayd in adrenal Right abrdominal CT.	MR in Right adrenal abdomiial glandp.
Left Adrenal gland	Left abdoimnal gland in adrenal CT.	abominal MR gland in adrenal Left.
Duodenum	in bdominal Duodeunm CT.	Duodenum in abdominal MR.
Bladder	dabdominal Bladdre in CT.	N/A
Prostate/uterus	in Prostate/uterus CT abdominal.	N/A

C.1. CAPR Meta-Prompt

As detailed in Section 3.1 of the main paper, CAPR employs Qwen3-VL-8B-Instruct to normalize and canonicalize varied text prompts. The following listing provides the complete, unified meta-prompt given to the LLM. This single meta-prompt was used across all adaptation directions to parse, correct, and contextually enrich all prompts into the canonical format: [Target] in [Anatomical Site] [Modality].

```
You are a meticulous assistant specializing in refining and standardizing biomedical image analysis prompts. Your user is preparing a list of text prompts for the BiomedParse vision foundation model, which requires precise, context-aware inputs.

Your task is to correct, refine, and normalize a list of "dirty" prompts I will provide.

Core Requirements
You must address two types of issues:

1. Linguistic Error Correction: You must find and fix all errors in each individual prompt, including but not limited to the following:
   - Spelling Mistakes
   - Letter Transposition
   - Missing Words/Characters
   - Repeated Words

2. Contextual Standardization: The BiomedParse model performs best when prompts are standardized. You must normalize every prompt to follow this strict format:

   [Target] in [Anatomical Site] [Modality]

To do this, you must first infer the global context (the default Modality and Site) from the entire list of prompts.

Step-by-Step Process
Before providing the final answer, you must perform the following internal reasoning steps:

Step 1: Analyze Global Context
Read the complete list of prompts I provide. Infer the most likely Modality and Anatomical Site that applies to the entire batch.

Step 2: State Your Context
I have analyzed the list. The inferred global context is:
Modality = [Your Inferred Modality]
Site = [Your Inferred Site]

Step 3: Iterate and Refine
Go through each prompt from the original list one by one.

- Analyze: The original prompt is: [original_prompt]
- Correct: The linguistic correction is: [corrected_prompt]
- Standardize: The prompt is missing context. Applying the global context, the final standardized prompt is: [standardized_prompt]
  (if the corrected prompt is liver and context is Abdomen CT, the final prompt is liver in abdomen CT)
```

```
- Exception: If a prompt already specifies its own context, like: tumor in brain MRI, respect that context and only correct its linguistic errors.

Final Output Format
Your final response to me MUST ONLY contain the list of fully corrected and standardized prompts.

- Do NOT include your internal Step-by-Step Thinking Process in the final output.
- Separate each standardized prompt with [SEP].

Example Output:
liver in abdomen CT[SEP]right kidney in abdomen CT[SEP]pancreas in abdomen CT[SEP]tumor in brain MRI

Task Starts Now
Here is the list of prompts for you to correct and normalize. Note the subprompt is also separated by [SEP]"""
```

C.2. Target-Specific Input Prompts

To facilitate full reproducibility of our extensive evaluations in Section 4, we provide the exact prompts used for each anatomical target. These prompts were fed into Tell2Adapt to generate the quantitative and qualitative results presented in the main paper. The prompts are organized by target in Table 10, Table 11, and Table 12.

Table 10. Prompts for abdominal targets in Tell2Adapt.

Target	**Prompt in MR→CT**	**Prompt in CT→MR**
Spleen	Spleen in abdominal CT.	Spleen in abdominal MR.
Right Kidney	Right kidney in abdominal CT.	Right kidney in abdominal MR.
Left Kidney	Left kidney in abdominal CT.	Left kidney in abdominal MR.
Gallbladder	Gallbladder in abdominal CT.	Gallbladder in abdominal MR.
Esophagus	Esophagus in abdominal CT.	Esophagus in abdominal MR.
Liver	Liver in abdominal CT.	Liver in abdominal MR.
Stomach	Stomach in abdominal CT.	Stomach in abdominal MR.
Aorta	Aorta in abdominal CT.	Aorta in abdominal MR.
Vena-Cava	Vena-cava in abdominal CT.	Vena-cava in abdominal MR.
Pancreas	Pancreas in abdominal CT.	Pancreas in abdominal MR.
Right Adrenal gland	Right adrenal gland in abdominal CT.	Right adrenal gland in abdominal MR.
Left Adrenal gland	Left adrenal gland in abdominal CT.	Left adrenal gland in abdominal MR.
Duodenum	Duodenum in abdominal CT.	Duodenum in abdominal MR.
Bladder	Bladder in abdominal CT.	N/A
Prostate/uterus	Prostate/uterus in abdominal CT.	N/A

Table 11. Prompts for brain targets in Tell2Adapt.

Target	**Prompt in T1n→T2w**	**Prompt in T2w→T1n**	**Prompt in T1c→T2f**	**Prompt in T2f→T1c**
TC	Non-enhancing tumor core in head T2 weighted MR.	Non-enhancing tumor core in head MR naive T1.	Non-enhancing tumor core in head MR T2 FLAIR.	Non-enhancing tumor core in head post-contrast T1 MR.
SNFH	Surrounding non-enhancing FLAIR hyperintensity in head T2 weighted MR	Surrounding non-enhancing FLAIR hyperintensity in head MR naive T1.	Surrounding non-enhancing FLAIR hyperintensity in head MR T2 FLAIR.	Surrounding non-enhancing FLAIR hyperintensity in head post-contrast T1 MR.
ET	Enhancing tissue in head T2 weighted MR.	Enhancing tissue in head MR naive T1.	Enhancing tissue in head MR T2 FLAIR.	Enhancing tissue in head post-contrast T1 MR.
RC	Resection cavity in head T2 weighted MR.	Resection cavity in head MR naive T1.	Resection cavity in head MR T2 FLAIR.	Resection cavity in head post-contrast T1 MR.

Table 12. Prompts for cardiac targets and polyp in Tell2Adapt.

Target	**Prompt in MR→US**	**Prompt in US→MR**
Left Ventricle	Left ventricle in heart ultrasound.	Left ventricle in cardiac MRI.
Myocardium	Myocardium in heart ultrasound.	Myocardium in cardiac MRI.
Target	**Prompt in Kvasir→CVCDB**	**Prompt in CVCDB→Kvasir**
Polyp	Polyp in colon endoscopes.	Polyp in colon endoscopes.