# Supplementary: Rethinking Annotator Simulation: Realistic Evaluation of Whole-Body PET Lesion Interactive Segmentation Methods

Zdravko Marinov[1,2], Moon Kim[3], Jens Kleesiek[3,4], and Rainer Stiefelhagen[1]

[1] Karlsruhe Institute of Technology, Karlsruhe, Germany
[2] HIDSS4Health - Helmholtz Information and Data Science School for Health, Karlsruhe/Heidelberg, Germany
[3] Institute for AI in Medicine, University Hospital Essen, Essen, Germany
[4] Cancer Research Center Cologne Essen (CCCE), University Medicine Essen, Essen,
[1]{firstname.lastname@kit.edu} , [3]{firstname.lastname@uk-essen.de}

## 1 Notation Table

Table 1: Notation table for all our equations.

| | |
|---|---|
| $R$ | Simulated robot user |
| $A$ | Real annotator |
| $\mathcal{A}$ | Set of all annotators in a user study |
| $\mathcal{I}$ | Set of all PET images in a user study |
| $\mathcal{U}_{[a,b]}$ | Uniform distribution over the interval $[a, b]$ |
| $\{a \mid a \in X \text{ and } \texttt{condition}(a)\}$ | Set of all elements belonging to $X$ that fulfill $\texttt{condition}(\cdot)$ |
| $W \times H \times D$ | Dimensions of the PET images (Width, Height, Depth) |
| $I \in \mathbb{R}^{W\times H\times D}$ | PET image |
| $I_Y \in \{0,1\}^{W\times H\times D}$ | Ground-truth label of image $I$ |
| $\texttt{clicks}(X, I)[i] \in \mathbb{N}^3$ | The $i^{\text{th}}$ click placed by $X \in \{R, A\}$ in image $I$ |
| $\texttt{pred}(I)[i] \in \{0,1\}^{W\times H\times D}$ | The model prediction after placing the $i^{\text{th}}$ click in image $I$ |
| $\texttt{err}(I)[i] \in \{0,1\}^{W\times H\times D}$ | Missegmentated regions in the model prediction $\texttt{pred}(I)[i]$ |
| $\texttt{center}(\cdot) : \{0,1\}^{W\times H\times D} \to \mathbb{N}^3$ | Geometric center, i.e., point that is furthest away from all boundaries |
| $\texttt{largest\_component}(\cdot) : \{0,1\}^{W\times H\times D} \to \{0,1\}^{W\times H\times D}$ | Largest connected component in a binary tensor |
| $\texttt{uniform}(X)$ | Uniform distribution over all non-zero entries in $X$ |
| $\texttt{epistemic}(X)$ | Distribution defined over the normalized epistemic uncertainty values |
| $\texttt{EDT}(X)$ | Normalized Euclidean Distance Transform over the non-zero entries in $X$ |
| $\texttt{SUV}(I, I_Y)$ | Distribution defined over the normalized SUV values in $I$, where $I_Y = 0$ |
| $p_{\text{system}} \in [0, 1]$ | Probability of systematic label non-conforming click in our robot user |
| $p_{\text{perturb}} \in [0, 1]$ | Probability of random spatial perturbation in our robot user |
| $a \in \mathbb{N}$ | Perturbation amplitude |
| $[\![X]\!]$ | Iverson Bracket $[\![X]\!] = 1$ if $X$ is true, otherwise $[\![X]\!] = 0$ |
| $\bar{C}(A, I)$ | Set of label conforming clicks of annotator $A$ for image $I$ |
| $\hat{C}(A, I)$ | Set of label non-conforming clicks of annotator $A$ for image $I$ |
| $\texttt{bound}(c, I_Y)$ | Minimum distance of click $c$ to the boundary of the label $I_Y$ |
| $\texttt{cent\_dist}(c, I_Y)$ | Minimum distance of click $c$ to the center of the label $I_Y$ |
| $\texttt{components}(X)$ | Set of all connected components of $X$ |
| $d(c, I_Y)$ | Minimum Euclidean distance of click $c$ to the ground-truth label $I_Y$ |
| $\texttt{dice}(A, I)[i]$ | Dice score after annotator $A$'s $i^{\text{th}}$ click on image $I$ |

## 2 Annotator Background

Table 2: Background of all annotators in our two user studies.

| | | | | |
|---|---|---|---|---|
| User Study 1 | Annotator 1<br>Medical Doctor | Annotator 2<br>Medical Student | Annotator 3<br>Radiologist | Annotator 4<br>Medical Student |
| User Study 2 | Annotator 1<br>Medical Doctor | Annotator 2<br>Medical Student | Annotator 3<br>Medical Student | Annotator 4<br>Radiologist |

## 3 Joint Probability Derivation

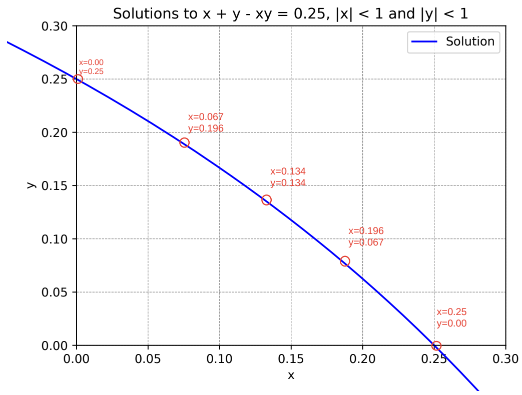


Fig. 1: Solutions to the joint probability of $\mathbb{P}(\texttt{perturb}(\cdot)\text{ OR }\texttt{system}(\cdot)) = 0.25$. As they are independent $\mathbb{P}(\texttt{perturb}(\cdot)\text{ OR }\texttt{system}(\cdot)) = p_{\text{perturb}} + p_{\text{system}} - p_{\text{perturb}} \cdot p_{\text{system}}$. We denote $p_{\text{system}}$ as $x$ and $p_{\text{perturb}}$ as $y$ in the plot and sample equidistant values from the solutions to obtain the final probabilities for our experiments.