A. Data Pre-processing

We make use of the publicly available anatomy-centered annotations provided by Chest ImaGemone [9, 10] to obtain masking maps of anatomical regions ($\{\mathcal{M}_n^{(\text{Abn})}\}_{n=1}^{N}, \mathcal{M}^{(\text{Chest})}$) and derive the hierarchical diagnostic description $\{H_m\}_{m=0}^{M}$ for each report in the MIMIC CXR dataset that is used for our performance evaluation.

The Chest ImaGenome dataset (https://physionet.org/content/chest-imagenome/1.0.0) contains scene graphs of each frontal chest X-ray image in the MIMIC CXR dataset. Each scene graph contains the bounding box coordinates of 29 anatomical regions presented in the image. The bounding box is associated with several sentences of the report if these sentences is about the clinical findings observed on that anatomical part. The abnormality terms mentioned in each sentence is also annotated.

A.1. Constructing Masking Maps of Anatomical Regions

To obtain the masking maps, we first collect the coordinates of the pre-defined N anatomical regions, denoted as $\{(\mathrm{x}_n, \mathrm{y}_n, \mathrm{w}_n, \mathrm{h}_n)\}_{n=1}^{N}$.

Consequently, given an X-ray image of the size of W × H, we propose to obtain a masking map $\mathcal{M}_n^{(\text{Ana})} \in \mathbb{R}^{\mathcal{W}\times\mathcal{H}}$ for each anatomical parts with $\{0, 1\}$ value indicating whether the bounding box $(\mathrm{x}_n, \mathrm{y}_n, \mathrm{w}_n, \mathrm{h}_n)$ of the detected anatomical part falls on the corresponding feature region of the size of $\mathcal{H}\times\mathcal{W}$. The implementation is described as follows:

i) $(\mathrm{x}_n, \mathrm{y}_n, \mathrm{w}_n, \mathrm{h}_n)$ the coordinates of each anatomical part are re-scaled from W × H to $\mathcal{W} \times \mathcal{H}$, denoted as $(\mathrm{x}_{1|n}, \mathrm{x}_{2|n}, \mathrm{y}_{1|n}, \mathrm{y}_{2|n})$;

ii) An all-zero matrix with size $\mathcal{M}_n^{(\text{abn})} \in \mathbb{R}^{\mathcal{W}\times\mathcal{H}}$ is initialized for each anatomical part, and the values of the matrix are set to 1, denoted as $\mathcal{M}_n^{(\text{Ana})}[i][j] = 1$, of which the row index i and the column index j subject to $i \in [x_{1|n}, x_{2|n}]$ and $j \in [y_{1|n}, y_{2|n}]$.

Next, we obtain a masking map of the overall chest area, denoted $\mathcal{M}^{(\text{Chest})}$. We start by constructing the coordinates of the overall chest area by collecting the minimum values of $\{x_{1|n}\}_{n=1}^{N}$ and $\{y_{1|n}\}_{n=1}^{N}$, and the maximum values of $\{x_{2|n}\}_{n=1}^{N}$ and $\{y_{2|n}\}_{n=1}^{N}$, to construct $(\mathrm{x}_1^{(\text{Chest})}, \mathrm{x}_2^{(\text{Chest})}, \mathrm{y}_1^{(\text{Chest})}, \mathrm{y}_2^{(\text{Chest})})$. Similarly to $\mathcal{M}_n^{(\text{Ana})}$, an all-zero matrix is created, denoted $\mathcal{M}^{(\text{Chest})} \in \mathbb{R}^{\mathcal{W}\times\mathcal{H}}$, where $\mathcal{M}^{(\text{Chest})}[i][j] = 1$ if i and j subject to $i \in [\mathrm{x}_1^{(\text{Chest})}, \mathrm{x}_2^{(\text{Chest})}]$ and $j \in [\mathrm{y}_1^{(\text{Chest})}, \mathrm{y}_2^{(\text{Chest})}]$. The pseudocode of the corresponding algorithm is shown in Alg. 1.

A.2. Constructing Hierarchical Diagnostic Descriptions

To obtain three-level hierarchical diagnostic descriptions $\{H_m\}_{m=0}^{M=2}$, at each level we construct the corresponding labels of the diagnostic items and then construct the diagnostic descriptions using a few curated templates with blanks filled with the constructed annotations.

To begin with, we make use of abnormality annotations provided by Chest ImaGemone as the core abnormality labels to construct the labels of diagnostic items at each level. To obtain abnormality labels $\{l_{i|(n,k)}^{(\text{Abn})}\}$, for each n^{th} anatomical part involved in the i^{th} report, we collect the associated annotations of the pre-defined K abnormalities, and then construct the labels of abnormality presentation, denoted as $\{l_{i|(n,k)}^{(\text{Abn})}\}_{k=1}^{K}$. $l_{i|(n,k)}^{(\text{Abn})} = 1$ indicates that the k^{th} abnormality is presented on the n^{th} anatomical parts; and $l_{i|(n,k)}^{(\text{Abn})} = 0$ otherwise.

Consequently, we construct the labels of diagnostic items at each level. The implementation details of the three-level diagnostic descriptions are presented as follows:

i) $H_0 \in \mathbb{R}^{N\times L}$ contains N diagnostic items which describe whether the size and location of the N anatomical parts are normal or not.
Implementation: We construct the labels of location normality $\{\{l_{i|n}^{(\text{loc})}\}_{n=1}^{N}\}$ and size normality $\{\{l_{i|n}^{(\text{size})}\}_{n=1}^{N}\}$ of the N anatomical parts by their bounding box coordinates of $\{\{(\mathrm{x}_{i|n}, \mathrm{y}_{i|n}, \mathrm{w}_{i|n}, \mathrm{h}_{i|n})\}_{n=1}^{N}\}$ as described in Alg. 2.
ii) $H_1 \in \mathbb{R}^{N\times L}$ contains N diagnostic items which describe whether the N anatomical part are normal and with some medical devices or not;
Implementation: For each i^{th} report, we construct the labels of anatomical normality $\{\{l_{i|n}^{(\text{nor})}\}_{n=1}^{N}\}$ and medical device $\{\{l_{i|n}^{(\text{dev})}\}_{n=1}^{N}\}$ of the N anatomical parts.

To obtain the labels of anatomical normality $l_{i|n}^{(\text{nor})}$ of n^{th} anatomical part, we make use of $\{l_{i|(n,k)}^{(\text{Abn})}\}_{k=1}^{K}$ to construct

$$l_{i|n}^{(\text{nor})} = \begin{cases} 0 & \exists k \in [1, K], l_{i|(n,k)}^{(\text{Abn})} = 1; \\ 1 & otherwise. \end{cases} \tag{1}$$

We also obtain $l_{i|n}^{(\text{dev})}$ by setting $l_{i|n}^{(\text{dev})} = 1$ if there exist the associated sentences of n^{th} anatomical part which describe the placement of any medical device, tubes or lines; and $l_{i|n}^{(\text{dev})} = 0$ otherwise.
iii) $H_2 \in \mathbb{R}^{NK\times L}$ contains NK diagnostic items indicating if a pre-defined set of K abnormalities are detected in N anatomical parts or not.
Implementation: We use $l_{i|(n,k)}^{(\text{abn})}$ directly as the labels of the diagnostic items at the level $m = 2$ for each report.

Algorithm 1: Construct masking maps of anatomical parts and chest area

```
Input: The bounding box coordinate sets $\{(\mathrm{x}_n, \mathrm{y}_n, \mathrm{w}_n, \mathrm{h}_n\}_{n=1}^{N}$, the height and width of the image $(\mathrm{H}, \mathrm{W})$, the
       height and width of visual feature map $(\mathcal{H}, \mathcal{W})$
Output: The masking map of anatomical parts $\{\mathcal{M}_n^{(\mathrm{Ana})}\}_{n=1}^{N}$ and chest area $\mathcal{M}^{\mathrm{Chest}}$
def ReScale($(\mathrm{x}, \mathrm{y}, \mathrm{w}, \mathrm{h}), (\mathrm{H}, \mathrm{W}), (\mathcal{H}, \mathcal{W})$):
    $\mathrm{x}_1 = \mathrm{x} * \mathcal{W}/\mathrm{W}$;
    $\mathrm{x}_2 = (\mathrm{x} + \mathrm{w}) * \mathcal{W}/\mathrm{W}$;
    $\mathrm{y}_1 = \mathrm{y} * \mathcal{H}/\mathrm{H}$;
    $\mathrm{y}_2 = (\mathrm{y} + \mathrm{h}) * \mathcal{H}/\mathrm{H}$;
    return $(\mathrm{x}_1, \mathrm{x}_2, \mathrm{y}_1, \mathrm{y}_2)$;
for each $n \in N$ do
    Initial all-zero matrix $\mathcal{M}_n^{(\mathrm{Ana})} \in \mathbb{R}^{\mathcal{W} \times \mathcal{H}}$;
    $(\mathrm{x}_{1|n}, \mathrm{x}_{2|n}, \mathrm{y}_{1|n}, \mathrm{y}_{2|n}) \leftarrow$ReScale($(\mathrm{x}_n, \mathrm{y}_n, \mathrm{w}_n, \mathrm{h}_n), (\mathrm{H}, \mathrm{W}), (\mathcal{H}, \mathcal{W})$);
    for each $i \in [\mathrm{x}_{1|n}, \mathrm{x}_{2|n}]$ do
        for each $j \in [\mathrm{y}_{1|n}, \mathrm{y}_{2|n}]$ do
            $\mathcal{M}_n^{(\mathrm{Ana})}[i][j] = 1$
$\mathrm{x}^{(\mathrm{Chest})}, \mathrm{y}^{(\mathrm{Chest})} = \min(\{\mathrm{x}_n\})_{n=1}^{N}, \min(\{\mathrm{y}_n\})_{n=1}^{N}$;
$\mathrm{w}^{(\mathrm{Chest})}, \mathrm{h}^{(\mathrm{Chest})} = \max(\{\mathrm{x}_n + \mathrm{w}_n - \mathrm{x}^{(\mathrm{Chest})}\})_{n=1}^{N}, \max(\{\mathrm{y}_n + \mathrm{h}_n - \mathrm{h}^{(\mathrm{Chest})}\})_{n=1}^{N}$;
Initial all-zero matrix $\mathcal{M}^{(\mathrm{Chest})} \in \mathbb{R}^{\mathcal{W} \times \mathcal{H}}$;
$(\mathrm{x}_1^{(\mathrm{Chest})}, \mathrm{x}_2^{(\mathrm{Chest})}, \mathrm{y}_1^{(\mathrm{Chest})}, \mathrm{y}_2^{(\mathrm{Chest})})$ =ReScale($(\mathrm{x}^{(\mathrm{Chest})}, \mathrm{y}^{(\mathrm{Chest})}, \mathrm{w}^{(\mathrm{Chest})}, \mathrm{h}^{(\mathrm{Chest})}), (\mathcal{H}, \mathcal{W})$);
for each $i \in [\mathrm{x}_1^{(\mathrm{Chest})}, \mathrm{x}_2^{(\mathrm{Chest})}]$ do
    for each $j \in [\mathrm{y}_1^{(\mathrm{Chest})}, \mathrm{y}_2^{(\mathrm{Chest})}]$ do
        $\mathcal{M}^{(\mathrm{Chest})}[i][j] = 1$
return $\{\mathcal{M}_n^{(\mathrm{Ana})}\}_{n=1}^{N}, \mathcal{M}^{(\mathrm{Chest})}$;
```

Hierarchical diagnostic descriptions $\{H_m\}$ are then constructed by the labels of diagnostic items and a set of predefined templates, as illustrated in Table. 1. We note that the diversity of the curated templates could affect the performance of the proposed model to be learned. However, using manual-design templates is one available way to construct hierarchical diagnostic descriptions. The intelligent and optimal construction of diagnostic descriptions would be systematically studied in future research.

B. Model Implementation

The proposed model is implemented by PyTorch [6]. with a Python Package *Transformer* (https://huggingface.co/) adopted for the CLIP model. It takes 20 hours and 31 hours to fine-tune the CLIP(SapBERT) and learn our proposed model AHIVE on one NVIDIA 80G A100 GPU, respectively. For inference, it takes on average 0.398s to retrieve one report from 1000 candidates using the same GPU server.

All baselines are implemented according to their official codes and pre-trained weights[1]. The source code will be released after the review process.

Evaluation metric The abnormality classes evaluated in the evaluation metric are detailed as follows:

CE(11) covers the top-11 most frequent abnormalities: *Lung opacity*, *Mass/Nodule*, *Airspace opacity*, *Pleural Effusion*, *Atelectasis*, Pneumothorax, *Medical device/tubelines*, *Edema*, *Consolidation*, *Enlarged cardiac silhouette*, *Lung lesion*.

CE(11/5) covers the top-11 abnormalities for each of the following 5 anatomical parts: *Left lung*, *Right lung*, *Left hilar structures*, *Right hilar structures* and *Cardiac silhouette*.

CE(13+NL) covers 13 abnormalities considered in [3]: *Enlarged cardiomediastinum*, *Cardiomegaly*, *Lesion*, *Lung opacity*, *Edema*, *Consolidation*, *Pneumonia*, *Atelectasis*, *Pneumothorax*, *Pleural effusion*, *Other*, *Fracture*, *Support*

[1] CLIP [7]:https://huggingface.co/openai/clip-vit-large-patch14, SapBERT [5]: https://huggingface.co/cambridgeltl/SapBERT-from-PubMedBERT-fulltext, CXR-RePaiR [2]: https://github.com/rajpurkarlab/CXR-RePaiR, MdeCLIP [8]:https://github.com/RyanWangZf/MedCLIP, BiomedCLIP [11]:https://huggingface.co/microsoft/BiomedCLIP-PubMedBERT_256-vit_base_patch16_224, X-REM [4]:https://github.com/rajpurkarlab/X-REM

Template	H_0: "(1) is (2) the normal location. (1) is (3) the normal size." H_1: "(1) is (4) abnormal. There is (5) medical device in (1)." H_2: "There is (6) (7) in (1)."
Terms to fill	(1)=the term of n^{th} anatomical part (2)="in" if $l_n^{(\text{loc})} = 0$ otherwise "out of" (3)="in" if $l_n^{(\text{size})} = 0$ otherwise "out of" (4)="in" if $l_n^{(\text{nor})} = 0$ otherwise "not in" (5)="no" if $l_n^{(\text{dev})} = 0$ otherwise keep blank (6)="no" if $l_{n,k}^{(\text{abn})} = 0$ otherwise keep blank (7)=the term of k^{th} abnormality on n^{th} anatomical part

Table 1. An example of templates with blanks for constructing the hierarchical diagnostic descriptions, where `(#ID)` represents the blank to be filled according to the labels of the diagnostic items.

devices, and NL stands for *Normality* [1].

References

[1] Zhihong Chen, Yan Song, Tsung-Hui Chang, and Xiang Wan. Generating radiology reports via memory-driven transformer. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, pages 1439–1449, 2020. 3

[2] Mark Endo, Rayan Krishnan, Viswesh Krishna, Andrew Y Ng, and Pranav Rajpurkar. Retrieval-based chest x-ray report generation using a pre-trained contrastive language-image model. In *Machine Learning for Health*, pages 209–219. PMLR, 2021. 2

[3] Jeremy Irvin, Pranav Rajpurkar, Michael Ko, Yifan Yu, Silviana Ciurea-Ilcus, Chris Chute, Henrik Marklund, Behzad Haghgoo, Robyn Ball, Katie Shpanskaya, et al. Chexpert: A large chest radiograph dataset with uncertainty labels and expert comparison. In *Proceedings of the Conference of Association for the Advance of Artificial Intelligence*, pages 590–597, 2019. 2

[4] Jaehwan Jeong, Katherine Tian, Andrew Li, Sina Hartung, Subathra Adithan, Fardad Behzadi, Juan Calle, David Osayande, Michael Pohlen, and Pranav Rajpurkar. Multimodal image-text matching improves retrieval-based chest x-ray report generation. In *Medical Imaging with Deep Learning*, 2023. 2

[5] Fangyu Liu, Ehsan Shareghi, Zaiqiao Meng, Marco Basaldella, and Nigel Collier. Self-alignment pretraining for biomedical entity representations. *arXiv preprint arXiv:2010.11784*, 2020. 2

[6] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, et al. Pytorch: An imperative style, high-performance deep learning library. *Advances in Neural Information Processing systems*, 32, 2019. 2

[7] Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, et al. Learning transferable visual models from natural language supervision. In *International Conference on Machine Learning*, pages 8748–8763. PMLR, 2021. 2

[8] Zifeng Wang, Zhenbang Wu, Dinesh Agarwal, and Jimeng Sun. Medclip: Contrastive learning from unpaired medical images and text. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, pages 3876–3887, 2022. 2

[9] Joy T Wu, Nkechinyere N Agu, Ismini Lourentzou, Arjun Sharma, Joseph A Paguio, Jasper S Yao, Edward C Dee, William Mitchell, Satyananda Kashyap, Andrea Giovannini, et al. Chest imagenome dataset for clinical reasoning. *arXiv preprint arXiv:2108.00316*, 2021. 1

[10] Joy T Wu, Nkechinyere N Agu, Ismini Lourentzou, Arjun Sharma, Joseph A Paguio, Jasper S Yao, Edward C Dee, William Mitchell, Satyananda Kashyap, Andrea Giovannini, et al. Chest imagenome dataset (version 1.0. 0). *PhysioNet*, 5:18, 2021. 1

[11] Sheng Zhang, Yanbo Xu, Naoto Usuyama, Jaspreet Bagga, Robert Tinn, Sam Preston, Rajesh Rao, Mu Wei, Naveen Valluri, Cliff Wong, et al. Large-scale domain-specific pretraining for biomedical vision-language processing. *arXiv preprint arXiv:2303.00915*, 2023. 2

Algorithm 2: Construct labels of location and size normality

```
Input: The set of anatomical bounding box coordinate $\mathcal{B} = \{\{(\mathrm{x}_{i|n}, \mathrm{y}_{i|n}, \mathrm{w}_{i|n}, \mathrm{h}_{i|n})\}_{n=1}^{N}\}$, the set of normality
        labels $L_{(\text{normality})} = \{l_i^{(\text{nor})}\}$
Output: The sets of location normality labels $L_{(\text{location})} = \{\{l_{i|n}^{(\text{loc})}\}_{n=1}^{N}\}$ and size normality labels
        $L_{(\text{size})} = \{\{l_{i|n}^{(\text{size})}\}_{n=1}^{N}\}$
def RelativeSize($\{(\mathrm{x}_n, \mathrm{y}_n, \mathrm{w}_n, \mathrm{h}_n)\}_{n=1}^{N}$):
    $\mathrm{x}^{(\text{Chest})}, \mathrm{y}^{(\text{Chest})} = \min(\{\mathrm{x}_n\})_{n=1}^{N}, \min(\{\mathrm{y}_n\})_{n=1}^{N}$;
    $\mathrm{w}^{(\text{Chest})}, \mathrm{h}^{(\text{Chest})} = \max(\{\mathrm{x}_n + \mathrm{w}_n - \mathrm{x}^{(\text{Chest})}\})_{n=1}^{N}, \max(\{\mathrm{y}_n + \mathrm{h}_n - \mathrm{h}^{(\text{Chest})}\})_{n=1}^{N}$;
    for each $n \in N$ do
        $\tilde{s}_n = \mathrm{w}_n \cdot \mathrm{h}_n / (\mathrm{w}^{(\text{Chest})} \cdot \mathrm{h}^{(\text{Chest})})$;
    return $(\mathrm{x}^{(\text{Chest})}, \mathrm{y}^{(\text{Chest})}, \mathrm{w}^{(\text{Chest})}, \mathrm{h}^{(\text{Chest})})$,$\{\tilde{s}_n\}_{n=1}^{N}$;
def RelativeCenter($\{(\mathrm{x}_n, \mathrm{y}_n, \mathrm{w}_n, \mathrm{h}_n)\}_{n=1}^{N}, (\mathrm{x}^{(\textit{Chest})}, \mathrm{y}^{(\textit{Chest})}, \mathrm{w}^{(\textit{Chest})}, \mathrm{h}^{(\textit{Chest})})$):
    for each $n \in N$ do
        $\tilde{x}_n = \mathrm{x}_n + \mathrm{w}_n/2 - (\mathrm{x}^{(\text{Chest})} + \mathrm{w}^{(\text{Chest})}/2)/\mathrm{w}^{(\text{Chest})}$;
        $\tilde{y}_n = \mathrm{y}_n + \mathrm{h}_n/2 - (\mathrm{y}^{(\text{Chest})} + \mathrm{h}^{(\text{Chest})}/2)/\mathrm{h}^{(\text{Chest})}$;
    return $\{(\tilde{x}_n, \tilde{y}_n)\}$;
for each $n \in N$ do
    Initial $X_n = \emptyset$, $Y_n = \emptyset$, $S_n = \emptyset$ ;
for each $\{(\mathrm{x}_{i|n}, \mathrm{y}_{i|n}, \mathrm{w}_{i|n}, \mathrm{h}_{i|n})\}_{n=1}^{N} \in \mathcal{B}$ do
    if $l_i^{(nor)} = 1$ then
        $(\mathrm{x}_i^{(\text{Chest})}, \mathrm{y}_i^{(\text{Chest})}, \mathrm{w}_i^{(\text{Chest})}, \mathrm{h}_i^{(\text{Chest})})$, $\{\tilde{s}_{i|n}\}_{n=1}^{N} \leftarrow$ RelativeSize($\{(\mathrm{x}_{i|n}, \mathrm{y}_{i|n}, \mathrm{w}_{i|n}, \mathrm{h}_{i|n})\}_{n=1}^{N}$);
        $\{(\tilde{\mathrm{x}}_{i|n}, \tilde{\mathrm{y}}_{i|n})\}_{n=1}^{N} \leftarrow$ RelativeCenter($(\mathrm{x}_i^{(\text{Chest})}, \mathrm{y}_i^{(\text{Chest})}, \mathrm{w}_i^{(\text{Chest})}, \mathrm{h}_i^{(\text{Chest})})$,$\{(\mathrm{x}_{i|n}, \mathrm{y}_{i|n}, \mathrm{w}_{i|n}, \mathrm{h}_{i|n})\}_{n=1}^{N}$);
        for each $n \in N$ do
            Add $\tilde{\mathrm{x}}_{i|n}$ to $X_n$, add $\tilde{\mathrm{y}}_{i|n}$ to $Y_n$, add $\tilde{\mathrm{s}}_{i|n}$ to $S_n$;
for each $n \in N$ do
    $\mu_{x|n}$,$\sigma_{x|n} \leftarrow$ mean-and-std($X_n$);
    $\mu_{y|n}$,$\sigma_{y|n} \leftarrow$ mean-and-std($Y_n$);
    $\mu_{s|n}$,$\sigma_{s|n} \leftarrow$ mean-and-std($S_n$);
Initial $L_{(\text{location})} = \emptyset$, $L_{(\text{location})} = \emptyset$;
for each $\{(\mathrm{x}_{i|n}, \mathrm{y}_{i|n}, \mathrm{w}_{i|n}, \mathrm{h}_{i|n})\}_{n=1}^{N} \in \mathcal{B}$ do
    $(\mathrm{x}_i^{(\text{Chest})}, \mathrm{y}_i^{(\text{Chest})}, \mathrm{w}_i^{(\text{Chest})}, \mathrm{h}_i^{(\text{Chest})})$, $\{\tilde{s}_{i|n}\}_{n=1}^{N} \leftarrow$ RelativeSize($\{(\mathrm{x}_{i|n}, \mathrm{y}_{i|n}, \mathrm{w}_{i|n}, \mathrm{h}_{i|n})\}_{n=1}^{N}$);
    $\{(\tilde{\mathrm{x}}_{i|n}, \tilde{\mathrm{y}}_{i|n})\}_{n=1}^{N} \leftarrow$ RelativeCenter($(\mathrm{x}_i^{(\text{Chest})}, \mathrm{y}_i^{(\text{Chest})}, \mathrm{w}_i^{(\text{Chest})}, \mathrm{h}_i^{(\text{Chest})})$,$\{(\mathrm{x}_{i|n}, \mathrm{y}_{i|n}, \mathrm{w}_{i|n}, \mathrm{h}_{i|n})\}_{n=1}^{N}$);
    for each $n \in N$ do
        Initialize $l_{i|n}^{(\text{loc})} = 0$ and $l_{i|n}^{(\text{size})} = 0$;
        if $\tilde{\mathrm{x}}_{i|n} \in (\mu_{x|n} - \sigma_{x|n}, \mu_{x|n} + \sigma_{x|n})$ and $\tilde{\mathrm{y}}_n \in (\mu_{y|n} - \sigma_{y|n}, \mu_{y|n} + \sigma_{y|n})$ then
            Set $l_{i|n}^{(\text{loc})} = 1$;
        if $\tilde{s}_n \in (\mu_{s|n} - \sigma_{s|n}, \mu_{s|n} + \sigma_{s|n})$ then
            Set $l_{i|n}^{(\text{size})} = 1$;
        Add $l_{i|n}^{(\text{loc})}$ to $L_{(\text{location})}$, $l_{i|n}^{(\text{size})}$ to $L_{(\text{Size})}$;
return $L_{(\text{location})}, L_{(\text{size})}$;
```