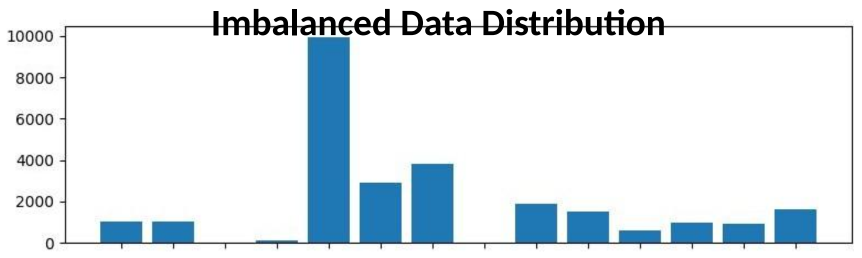
Multi-Modal Medical Image Augmentation for Controlled Heterogeneity and Fair Outcomes

Soo Yong Kim* [1,2], Seunghyeok Hong [2]

Seoul National University, Interdisciplinary Program in Artificial Intelligence, Seoul, Republic of Korea [1]
MODULABS [2]

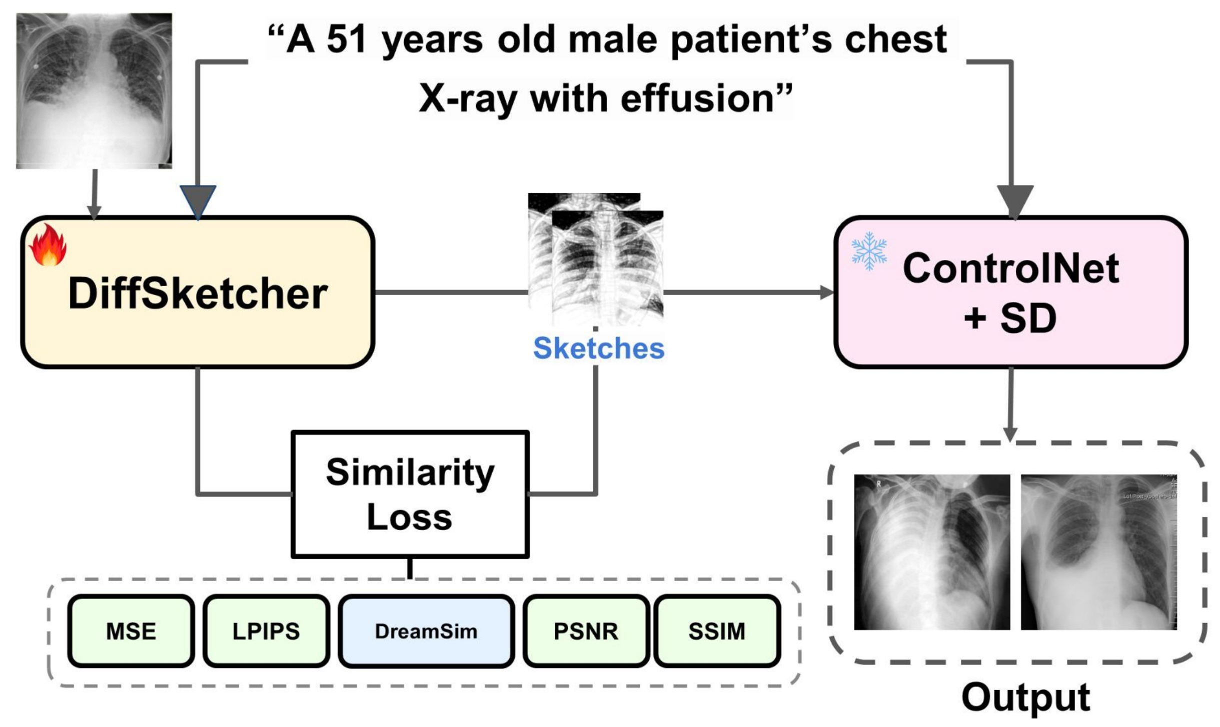
Motivation

- In medical domain, data imbalance among different patient group is critical problem.
- Data augmentation is a widely used strategy to solve this problem.
- However, generating specific patient groups data with various conditions (age, sex, disease) is challenging.



Our study aims to augment medical images under specific conditions

Abstract



Framework

(1) Select patient groups for data augmentation

: we propose a new metric to measure necessity of data augmentation of patient groups

$$M_K = \frac{1}{2}\left(\frac{|D_K|}{|D|} + S_K\right) \quad (1) \qquad S_k = \frac{2 * \sum_{1\le i\le j\le |D_K|} d_{ij}}{|D_K|(|D_K|-1)} \quad (2)$$

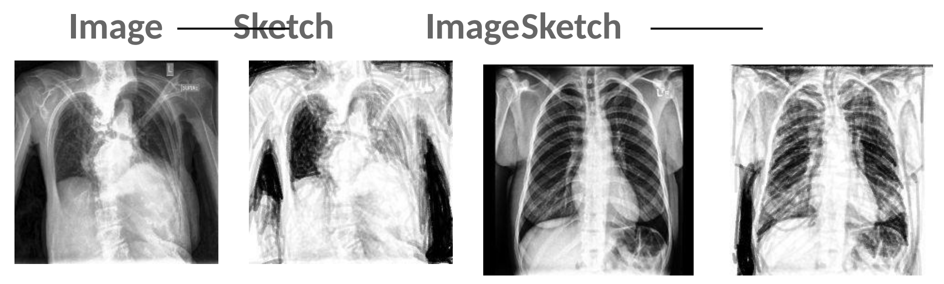
$$d_{ij} = \frac{1}{3} * \left[\frac{1}{C}\sum_{c=1}^{C} f(v_c^i, v_c^j) + \frac{|g^i//10 - g^j//10|}{g^m//10} + \delta(s^i, s^j)\right] \quad (3)$$

$$f(v_c^i, v_c^j) = \begin{cases} 1 & \text{if } v_c^i \neq v_c^j \\ 0 & otherwise \end{cases} \quad (4) \qquad \delta(s^i, s^j) = \begin{cases} 1 & \text{if } s^i \neq s^j \\ 0 & otherwise \end{cases} \quad (5)$$

M_K is the majority score and K is the index of patient groups. $|D|$ is the total number of data and $|D_K|$ is the number of data from K'th patient group. v is the label vector and c is the class index. g indicates the age of the patient and m is the index of the maximum age. s represents the sex of the patient.

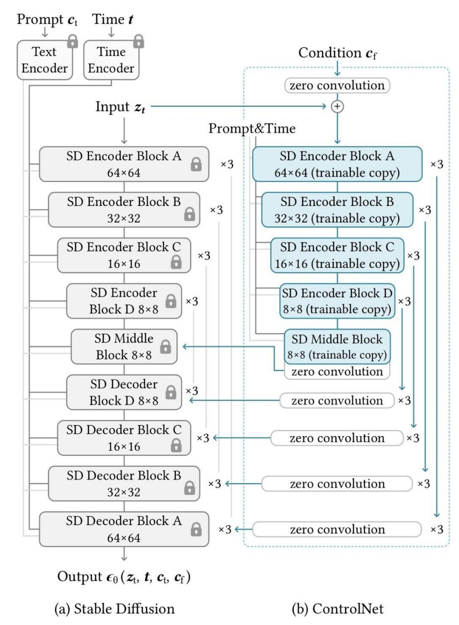
(2) Collect sketch-image pair using DiffSketcher

: DiffSketcher is vector graphics sketch generator. We pairs to train ControlNet.



(3) Fine-tune pre-trained ControlNet

: Below results are obtained from the training process.

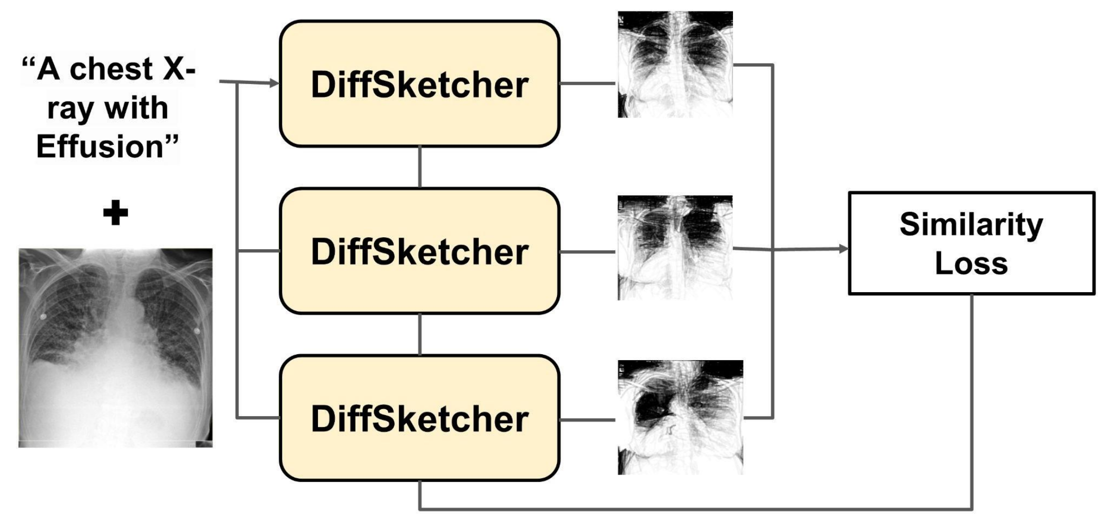


"A 70 years old female patient's chest x-ray with edema"

"A 48 years old male patient's chest x-ray with Consolidation, Atelectasis"

(4) Obtain diverse sketch data from DiffSketcher

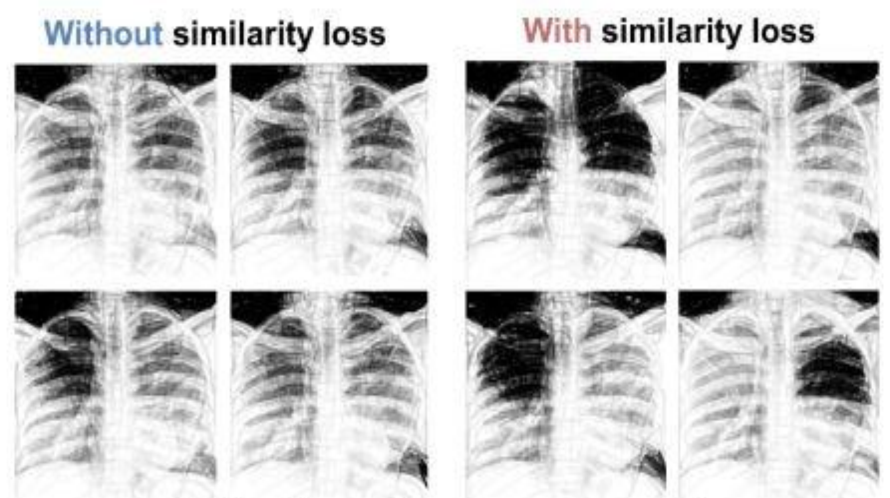
: we obtain multiple various sketched from DiffSketcher with the similarity loss which we propose.



Similarity loss: consists of both low-level and high-level distance metrics

$$\mathcal{L}_{sim}^{i} = \alpha * \frac{\sum_{1\le j\le n, i\neq j}(1 - h(x_i, x_j))}{n-1} + (1-\alpha) * \frac{\sum_{1\le j\le n, i\neq j}(1 - l(x_i, x_j))}{n-1} \quad (6)$$

α regulates the contributions of high-level and low-level similarity metrics. i and j are the index of the DiffSketcher. x is the input image. h represents high-level similarity metric and l is the low-level similarity metric. n is the number of sketches we generate at the same time.



"A chest X-ray with severe pneumonia"

(5) Measure diversity of augmented patient group

: we propose a new metric to measure diversity of dataset using saliency map.

$$D-score = \frac{2}{N(N-1)} \sum_{1\le i\le j\le N} \frac{1}{C} \sum_{c=1}^{C} \omega(A_c^i, A_c^j) \quad (7)$$

ω is the Wasserstein distance and A indicates saliency map. c represents the index of the convolutional block and N is the total number of generated samples.

IPAI 서울대학교 대학원 협동과정 인공지능전공 Interdisciplinary Program in Artificial Intelligence

Email: ksyint1111@snu.ac.kr (Soo Yong Kim) , shongdr@gmail.com (Seunghyeok Hong)