TB-SMGAN : A GAN Based Hybrid Data Augmentation Framework on Chest X-ray Images and Reports

Hasan Berat ÖZFİDAN, Mehmet Ulvi ŞİMŞEK*

ASELSAN, Ankara, Türkiye

M.S. Student in Software Convergence, Graduate School
Younguk Park
2025.05.13

SCH 순천향대학교 SOON CHUN HYANG UNIVERSITY

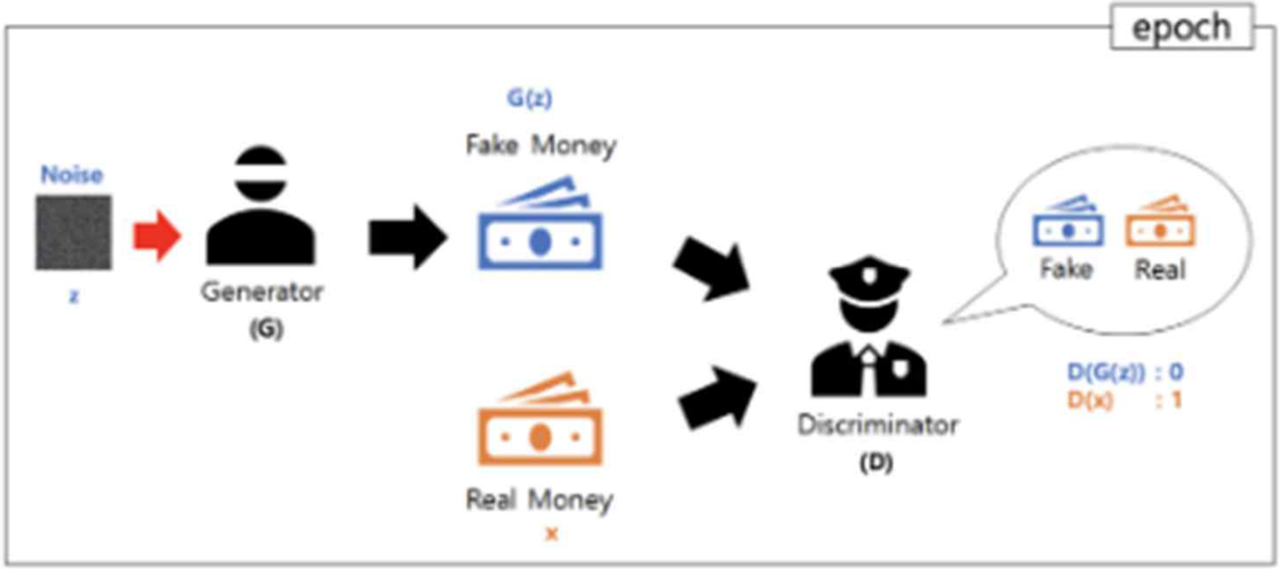
AI for Medical Imaging

Reason for paper selection

- GAN ? ->
- 이미지 생성
- 노이즈 제거/ 해상도 복원
- 프라이버시 보호용 데이터 생성
- 딥페이크 / 영상 생성 …
- 데이터 증강, 멀티모달

GAN?

◆ GAN(Generative Adversarial Networks, 생성적 적대 신경망)이란 비지도학습에 사용되는 머신러닝 프레임워크의 한 종류로 GAN은 다른 알고리즘과 달리 이전에는 없던 새로운 데이터를 생성 할 수 있음.



$$\min_{G}\max_{D} V(D,G) = \underbrace{\boxed{\mathbb{E}_{x\sim Pdata(x)}[\log D(x)]}}_{\text{진짜 데이터}} + \underbrace{\boxed{\mathbb{E}_{z\sim Pz(z)}[\log(1-D(G(z)))]}}_{\text{가짜 데이터}}$$

GAN의 손실함수

Introduction

- Accurately classifying medical images is crucial for early diagnosis and effective treatment, but limited training data remains a critical obstacle.
- Traditional data augmentation techniques like geometric transformations and color jittering are used to expand datasets in literature.
- Another strategy is Self-Supervised Learning, which holds promise by learning representations from unlabeled data. Despite its potential, aligning learned features with specific diagnosis tasks proves to be challenging, and the performance in medical domains often falls short of optimal levels.
- Transfer Learning, where leveraging pre-trained models on large-scale datasets can enhance performance on smaller medical datasets. However, the presence of domain mismatch, characterized by differences between the training and target datasets, poses a substantial challenge and can significantly impact the model's generalizability.
- This paper introduces a novel Text-Based Style-Manipulated GAN augmentation framework (TB-SMGAN) that overcomes these limitations by leveraging the combined capabilities of StyleGAN2-ADA (Karras et al., 2020) and StyleCLIP (Patashnik et al., 2021).

Introduction

- Rule-Based Information Extraction from X-ray Reports: We develop a novel rule-based algorithm for accurately extracting relevant information from X-ray reports. This extracted information provides additional context for medical image analysis and enhances the effectiveness of downstream tasks.

- Fine-Tuning CLIP for Medical Domain Adaptation: We fine-tune CLIP using various text extraction methods specifically tailored for the medical domain. This domain adaptation ensures the learned representations are relevant and informative for medical applications, leading to improved performance in subsequent tasks.

- Text-Based Latent Space Manipulations for Medical Data Augmentation: This work introduces a novel approach that utilizes text-based information to manipulate the latent space of GANs. This approach enables the generation of synthetic medical data that is not only realistic but also semantically aligned with the extracted textual information, further enriching the training dataset and improving the generalizability of deep learning models trained on augmented data.

Related Work

- Frid-Adar et adl, 2018
  - DCGAN, ACGAN을 사용해 3개 질환에 대해 이미지 생성
  - Real data + Fake data => classification accuracy 올라감.
- Kora Venu & Ravula, 2020
  - 폐렴 진단 모델에서 기존 증강 vs GAN증강 비교
- Deepshikha & Naman, 2020
  - 클래스 불균형때문에 정확도가 낮은 걸 GAN을 이용해 부족한 불균형 보완.
- Sundaram & Hulkund, 2021
  - GAN이 기존 증강보다 소수 클래스 성능 향상에 더 효과적

TB-SMGAN FrameWork

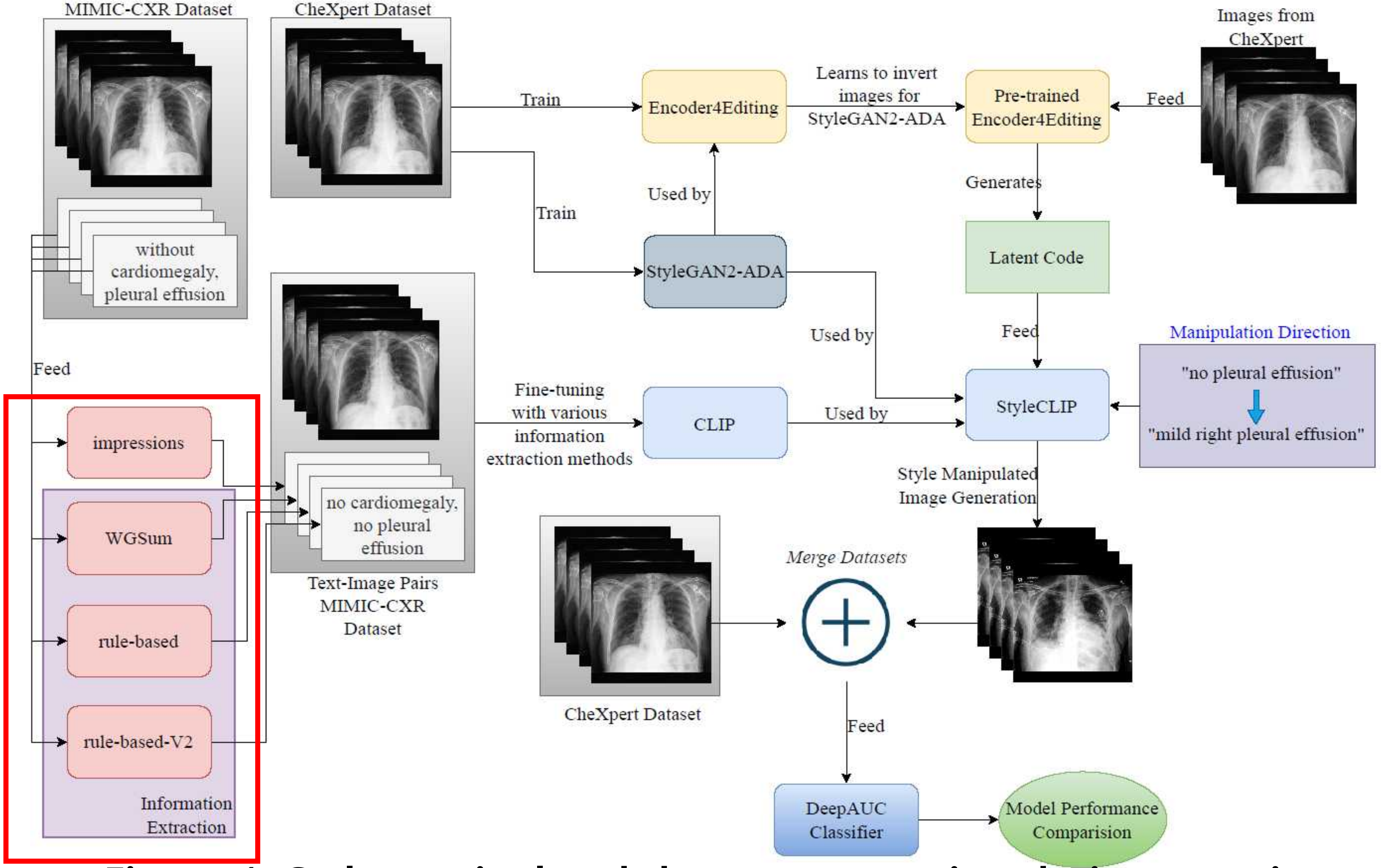


Figure 1. Style manipulated data augmentation design overview

TB-SMGAN FrameWork

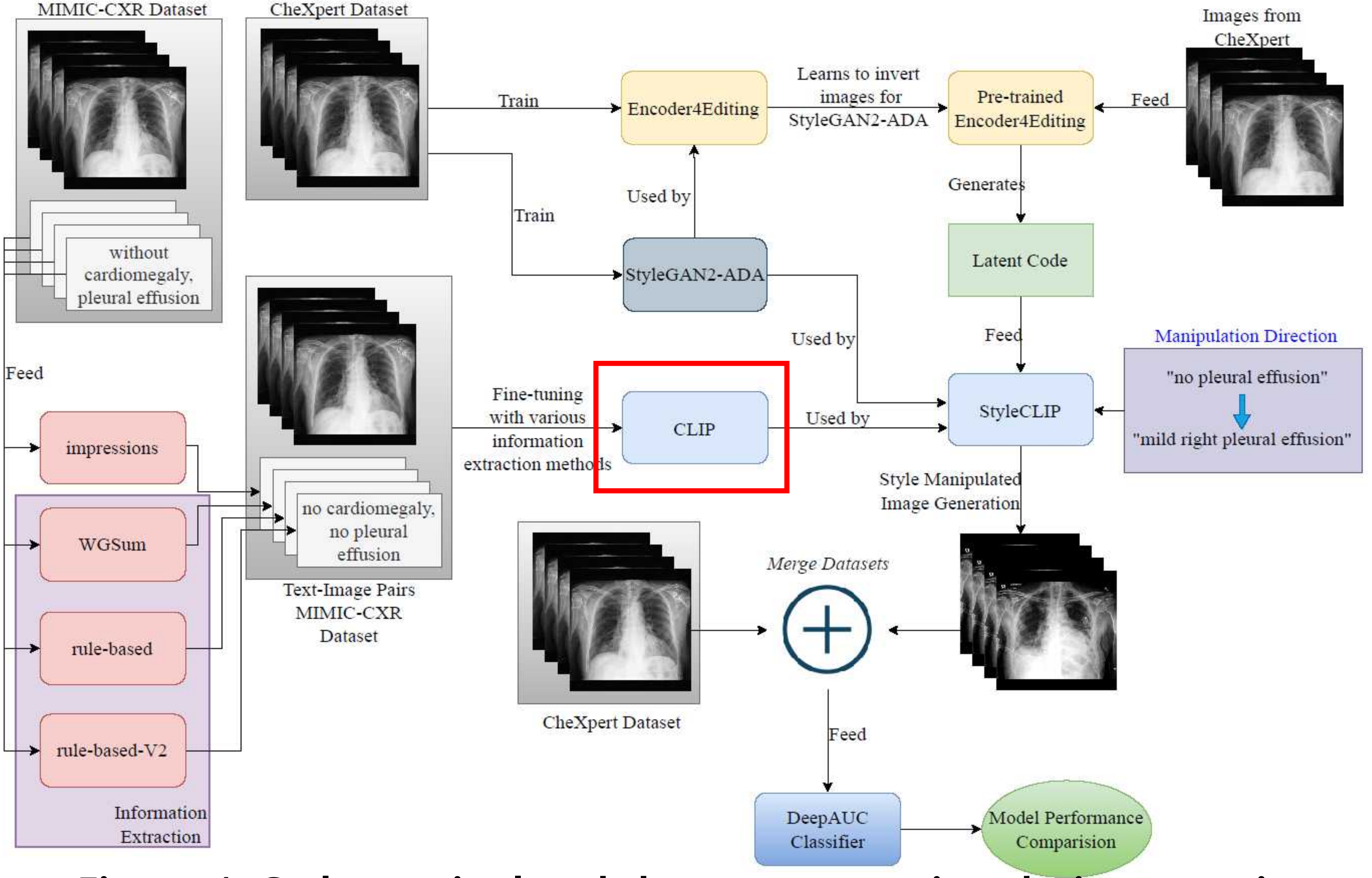


Figure 1. Style manipulated data augmentation design overview

AI for Medical Imaging

TB-SMGAN FrameWork

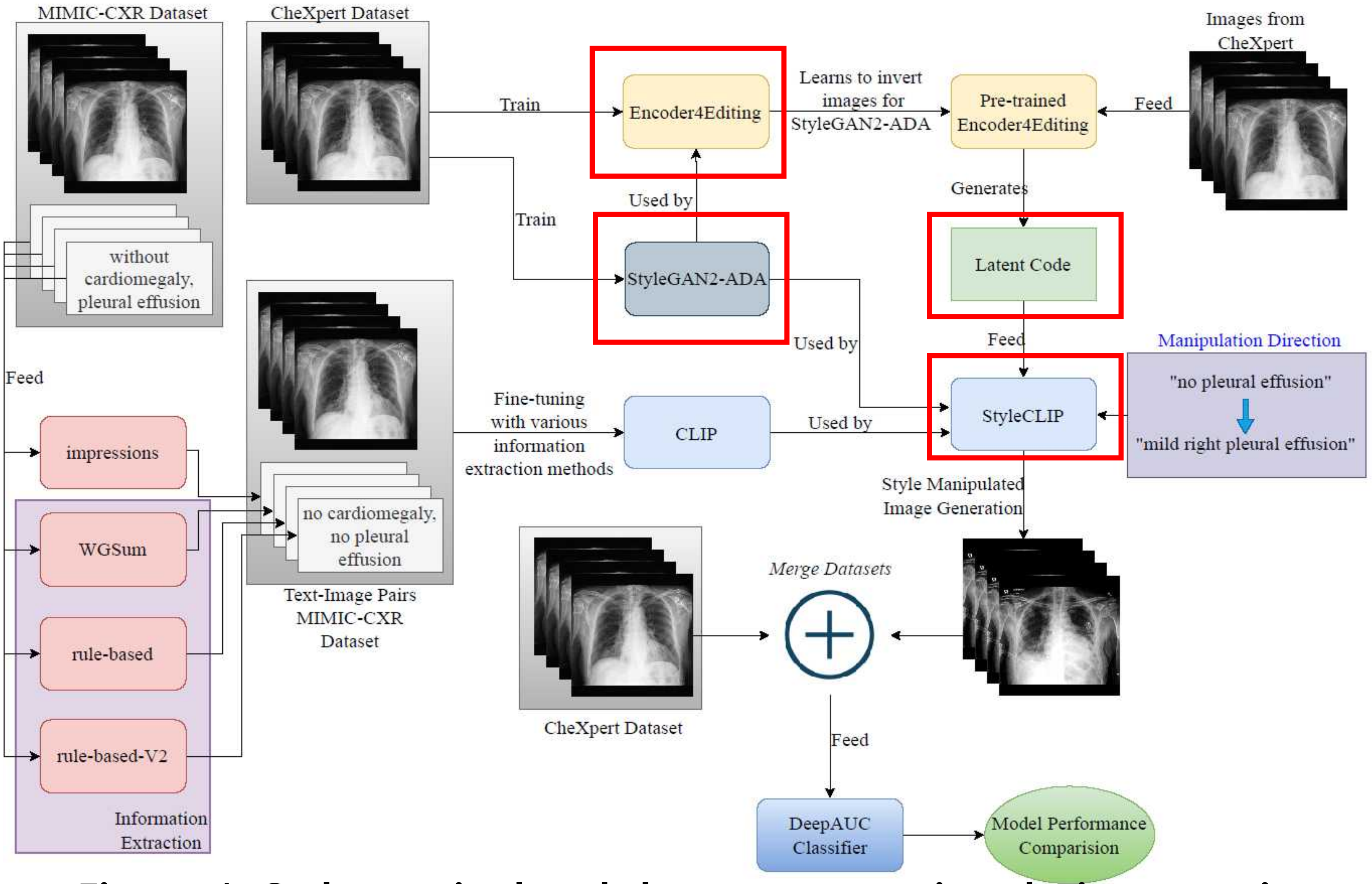


Figure 1. Style manipulated data augmentation design overview

AI for Medical Imaging

TB-SMGAN FrameWork

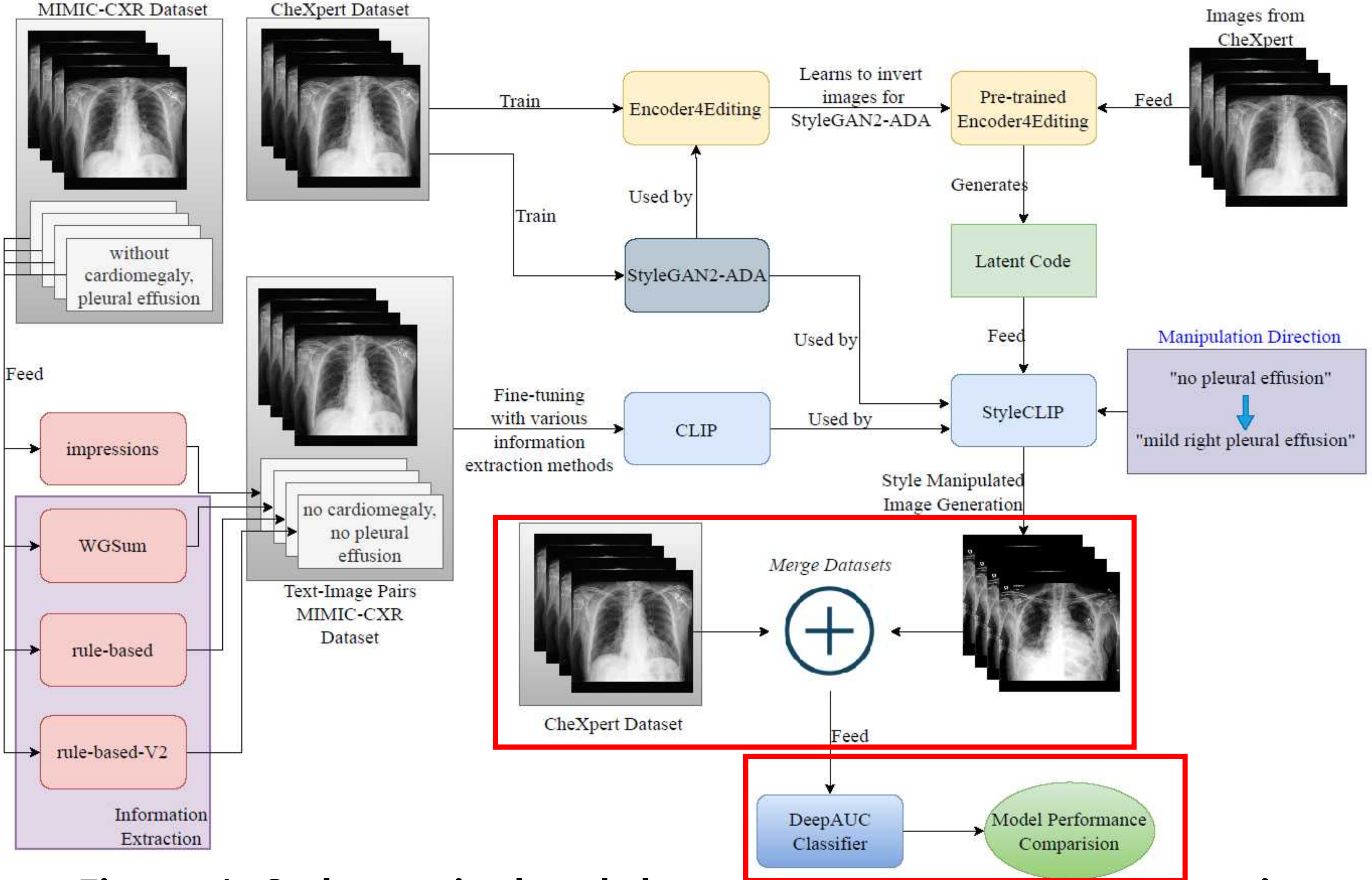


Figure 1. Style manipulated data augmentation design overview

AI for Medical Imaging

Experiment results

- **Datasets**
  - we employ the CheXpert dataset (Irvin et al., 2019), comprising about 225,000 chest X-ray images from Stanford University Medical Center.
  - The dataset includes 14 classes, with uncertain labels categorized as positive, negative, or uncertain.
  - On our framework, we focus on five specific classes—Atelectasis, Cardiomegaly, Consolidation, Edema, and Pleural Effusion.

Experiment results

- **TB-SMGAN Framework**
  - This section presents the experimental results of our proposed method for augmenting chest X-ray datasets using text-based style manipulations

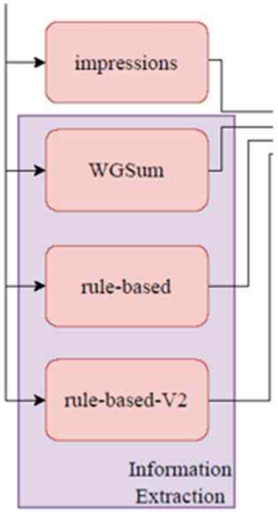


Table 1. *Classification results on only synthetic data generated by TB-SMGAN*

Fine-Tuning Strategy	Cardiomegaly (PR-AUC)	Edema (PR-AUC)	Consolidation (PR-AUC)	Atelectasis (PR-AUC)	Pleural-Effusion (PR-AUC)	Mean AUC
rule-based (-)	0,147	0,198	0,196	0,194	0,193	0,186
impression (-)	0,191	0,203	0,159	0,169	0,175	0,179
rule-based-V2 (-)	0,234	0,228	0,126	0,169	0,227	0,197
original-CLIP (-)	0,151	0,41	0,225	0,197	0,411	0,279
WGSum-generated (-)	0,254	0,241	0,207	0,247	0,24	0,238
randomized (-)	0,19	0,269	0,182	0,261	0,312	0,243
inverted	0,241	0,294	0,219	0,221	0,277	0,251
rule-based (+)	0,579	0,235	0,17	0,364	0,432	0,356
impression (+)	0,27	0,232	0,353	0,516	0,455	0,365
rule-based-V2 (+)	0,208	0,382	0,222	0,281	0,619	0,342
original-CLIP (+)	0,339	0,38	0,336	0,26	0,205	0,304
WGSum-generated (+)	0,159	0,579	0,281	0,221	0,315	0,311
randomized (+)	0,318	0,181	0,288	0,217	0,187	0,238

Experiment results

- **TB-SMGAN Framework**

Table 2. *Classification performance results of fine-tuning strategies on text-based style manipulated GAN augmented dataset*

Fine-Tuning Strategy	Cardiomegaly (PR-AUC)	Edema (PR-AUC)	Consolidation (PR-AUC)	Atelectasis (PR-AUC)	Pleural-Effusion (PR-AUC)	Mean AUC
pure-dataset	0,719	0,745	0,615	0,633	0,842	0,711
rule-based (+), inverted and rule-based-V2 (+)	0,707	0,756	0,592	0,644	0,847	0,709
impression (+)	0,688	0,766	0,562	0,68	0,855	0,71
original-clip (+)	**0,738**	0,779	0,566	0,667	0,843	0,719
rule-based-V2 (+)	0,658	0,8	0,643	0,66	0,841	0,72
rule-based (-)	0,641	0,782	0,728	0,624	0,837	0,722
inverted	0,7	0,788	0,66	0,651	0,841	0,728
rule-based (+)	0,684	0,796	0,659	0,693	0,821	0,731
WGSum-generated (-)	0,659	0,791	0,671	0,698	0,834	0,731
randomized (+)	0,678	**0,818**	0,614	0,712	0,835	0,731
rule-based-V2 (-)	0,649	0,787	0,703	0,707	0,838	0,737
randomized (-)	0,682	0,791	0,688	0,701	0,827	0,738
WGSum-generated (+)	0,687	0,79	0,669	0,691	0,852	0,738
ensemble-without-impression	0,694	0,796	0,69	0,701	0,818	0,74
impression (-)	0,664	0,796	0,758	0,682	0,823	0,744
StyleGAN2-Augmentation	0,71	0,772	0,676	0,702	**0,863**	0,745
original-clip (-)	0,67	0,81	0,706	0,727	0,827	0,748
rule-based (+) and inverted	0,723	0,796	0,658	**0,743**	0,847	0,753
TB-SMGAN (ensemble)	0,672	0,802	**0,783**	0,667	0,845	0,754

Conclusion

- This work leverages the generative power of StyleGAN2-ADA to perform data augmentation for medical image datasets.
- Moreover, we introduce a text-based style manipulated GAN augmentation technique named TB-SMGAN for the medical domain.
- We utilize DeepAUC, the top solution of the CheXpert competition, to demonstrate the effectiveness of our GAN augmentation technique.
- While our findings are promising, there are several directions for future research.
  - First, it would be beneficial to explore the application of our proposed framework to other domains beyond medical imaging, to assess its generalizability.
  - Second, developing a consistent measurement methodology for the quality of the dataset could yield valuable insights

Q&A