

# **BAYESIAN SELF-CRITICAL SEQUENCE TRAINING FOR MEDICAL REPORT GENERATION**

**Nguyễn Nguyễn Khang - 22520623**

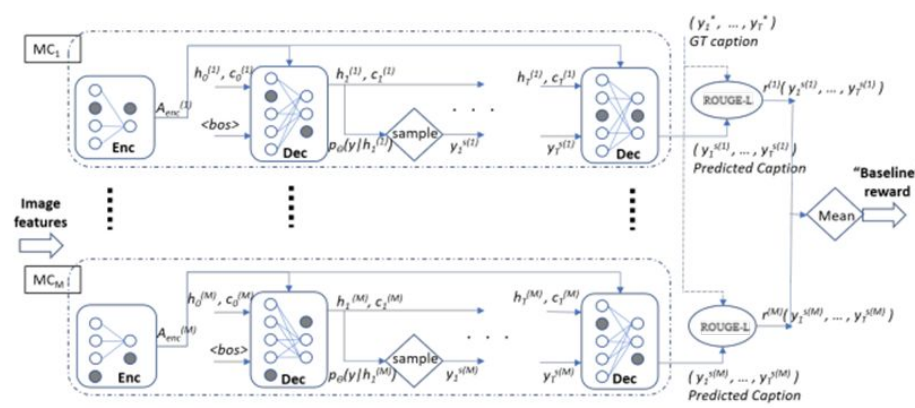
# Tóm tắt

- Lớp: CS519.021.KHTN
- Link Github:  
<https://github.com/KanzNguyen/CS519.021.KHTN>
- Link YouTube video:  
<https://www.youtube.com/watch?v=UGpcok5-vmo>
- Ảnh + Họ và Tên:

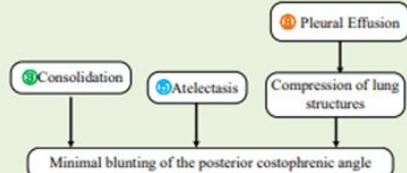


Nguyễn Nguyên Khang

# Giới thiệu



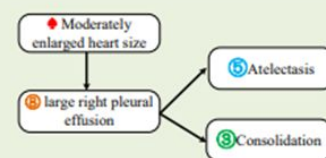
(a) ⑤③⑧



**Ground Truth:** Chest PA and lateral radiograph demonstrates a markedly elevated right hemidiaphragm with adjacent compressive ⑤atelectasis or ③consolidation. Minimal blunting of the posterior costophrenic angle ⑧may indicate a small right pleural effusion. Left lung is clear. ♦Cardiomediastinal borders are unremarkable.

**VLCI:** frontal and lateral views of the chest were obtained. ⑧there is a small right pleural effusion ⑤with overlying atelectasis ③underlying consolidation is not excluded. the left lung is clear. ♦the cardiac silhouette is not enlarged. the aorta is calcified and tortuous. ⑦no pneumothorax is seen.

(b) ③⑤⑧♦



**Ground Truth:** ♦The heart size is moderately enlarged. ♦The mediastinal silhouette and hilar contours are unchanged. A moderate to ⑧large right-sided pleural effusion is slightly increased in volume compared to prior examination with collapse of much of the right lower lobe and right middle lobe. ⑤There is also some consolidation at the base of the right upper lobe which could be due to ⑤compressive atelectasis. ⑥There is no left effusion. ⑦There is no pneumothorax

**VLCI:** there are low lung volumes which the bronchovascular markings. ⑧there is a large right pleural effusion with ⑤overlying atelectasis. ③underlying consolidation is not excluded. ⑦there is no pneumothorax. ♦the enlarged cardiac silhouette is due to the large right pleural effusion. mediastinal contours are grossly stable. no overt pulmonary edema is seen.

# Mục tiêu

- Nghiên cứu hiệu suất của mô hình VLCI [1] và phiên bản cải tiến VLCI with B-SCST trên các độ đo ROUGE-L, CIDER, BLEU, METEOR và CE
- Nghiên cứu hiệu suất của 2 mô hình trên 2 bộ dữ liệu IU-Xray và MIMIC-CXR sử dụng cách tiếp cận tăng cường dữ liệu truyền thống và sử dụng mô hình GAN.
- Phát triển một chương trình ứng dụng để minh họa hiệu năng của mô hình cải tiến.

# Nội dung

- **Tìm hiểu câu trả lời cho câu hỏi:** *“Liệu việc ứng dụng thuật toán B-SCST có khả năng huấn luyện nâng cao từ mô hình nền tảng được áp dụng thành công cho các bài toán image captioning, speech recognition và question answering [3][4][5] vào một bài toán hoàn toàn mới phức tạp hơn: sinh báo cáo y khoa từ ảnh chụp xét nghiệm có mang lại hiệu quả hơn so với kiến trúc ban đầu không?”* và câu hỏi *“Liệu cách tiếp cận mới trong bài toán tăng cường dữ liệu sử dụng các mô hình GAN trên tập dữ liệu huấn luyện có mang lại hiệu năng vượt trội hơn so với cách tiếp cận truyền thống hay không?”*.
- **Nghiên cứu hiệu quả của mô hình VLCL with B-SCST** cho bài toán sinh báo cáo y khoa (MRG) so với hiệu quả của mô hình ban đầu.
- **Nghiên cứu cách ứng dụng các mô hình GAN phổ biến** cho bài toán tăng cường dữ liệu hình ảnh y khoa bao gồm GAN truyền thống, DCGAN và CycleGAN [7][8][9].
- **Huấn luyện mô hình VLCL-base và VLCL with B-SCST trên 2 bộ dữ liệu IU-Xray và MIMIC-CXR**, so sánh hiệu quả của các chiến lược tăng cường dữ liệu theo cách truyền thống và cách sử dụng mô hình GAN.
- **Xây dựng chương trình ứng dụng.**

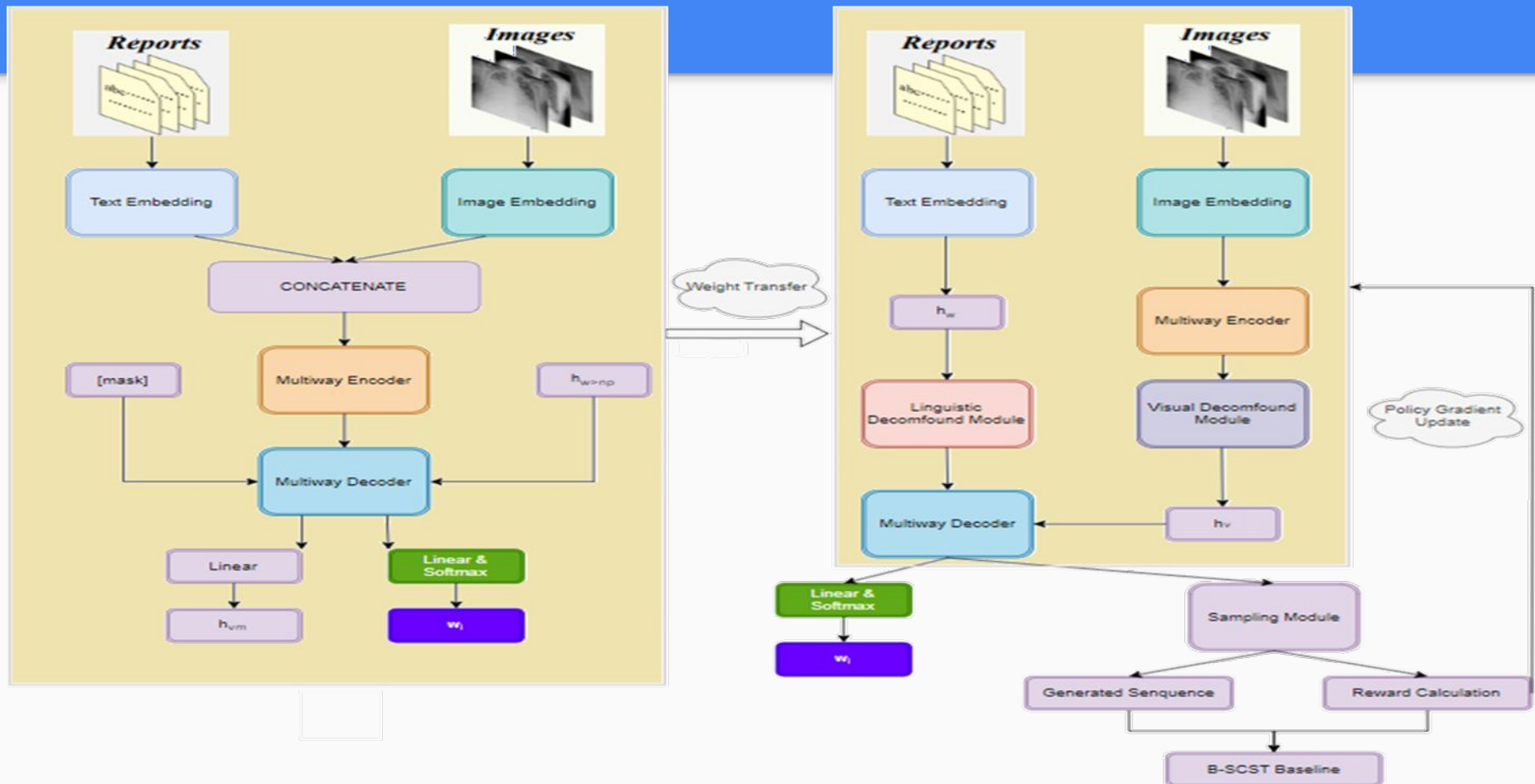
# Phương pháp

• **Nghiên cứu kiến trúc và quy trình vận hành của mô hình VLCI** có khả năng giảm thiểu ảnh hưởng hoặc thậm chí loại bỏ những thiên kiến thị giác và ngôn ngữ gây ra những mối quan hệ tương quan nhân quả không chính xác (do các confounders gây ra) trong quá trình huấn luyện.

• **Nghiên cứu cách áp dụng thuật toán B-SCST**, kỹ thuật này dựa trên thuật toán REINFORCE và SCST để tiến hành huấn luyện nâng cao cho một mô hình đã được huấn luyện trước đó để đồng nhất cách đánh giá hiệu năng mô hình trong quá trình huấn luyện và kiểm thử sử dụng các độ đo tương ứng là các hàm bất khả vi. Mô hình VLCI sau khi đã được huấn luyện sẽ được nối với một sampling module hoạt động theo nguyên tắc của B-SCST để tạo ra các chuỗi con và tính toán baseline, các chuỗi này sẽ được so sánh với baseline và dùng kết quả so sánh này để cập nhật tham số của mô hình dựa trên hàm mất mát được quy định bởi B-SCST.

• **Phân chia 2 tập dữ liệu thành các tập huấn luyện và kiểm thử**, tạo ra các phiên bản khác nhau trên tập huấn luyện sử dụng phương pháp tăng cường dữ liệu truyền thống, huấn luyện các mô hình GAN để tạo ra phiên bản dữ liệu huấn luyện có tăng cường dữ liệu sử dụng các lớp mô hình GAN.

• **Nghiên cứu cách đánh giá mô hình** tạo báo cáo y khoa trên các độ đo ROUGE-L, CIDER, BLEU, METEOR và CE.



# Phương pháp

- **Huấn luyện mô hình VLCl-base** trên 2 bộ dữ liệu IU-Xray và MIMIC-CXR trong cả trường hợp sử dụng phương pháp tăng cường dữ liệu truyền thống và phương pháp sử dụng các mô hình GAN.
- **Huấn luyện mô hình VLCl with B-SCST** từ mô hình VLCl-base đã được huấn luyện trước đó trên các phiên bản áp dụng tăng cường dữ liệu khác nhau của 2 bộ dữ liệu IU-Xray và MIMIC-CXR, hàm điểm thưởng sử dụng độ đo ROUGE-L đo xâu con chung dài nhất giữa chuỗi dự đoán và chuỗi tham chiếu, nhằm khuyến khích mô hình tiến hành cập nhật tham số để đạt được kết quả cao hơn trên độ đo này tức tạo ra chuỗi dự đoán khớp với chuỗi tham chiếu, đồng thời đồng nhất quá trình huấn luyện và kiểm thử.
- **Đánh giá hiệu năng của mô hình VLCl-with B-SCST so với mô hình VLCl-base** trên 2 bộ dữ liệu, trong cả trường hợp tăng cường dữ liệu theo hướng truyền thống và sử dụng GAN.
- Xây dựng chương trình ứng dụng để khảo sát hiệu năng của mô hình trên các ảnh chụp X-quang ngực để tạo ra các báo cáo y khoa tương ứng.



# Kết quả dự kiến

- Kết quả thực nghiệm của mô hình VLCL-with B-SCST mà chúng tôi phát triển ứng dụng trong bài toán tạo báo cáo y khoa, so sánh trên các độ đo ROUGE-L, CIDER, BLEU, CE với mô hình gốc, trên hai bộ dữ liệu với các phiên bản sử dụng tăng cường dữ liệu bằng các tiếp cận truyền thống và các mô hình GAN.
- Một chương trình ứng dụng để minh họa hiệu năng của mô hình.

# Tài liệu tham khảo

- [1] Cross-Modal Causal Intervention for Medical Report Generation Weixing Chen, Yang Liu, Member, IEEE, Ce Wang, Jiarui Zhu, Shen Zhao, Guanbin Li, Member, IEEE, Cheng-Lin Liu, Fellow, IEEE, and Liang Lin, Fellow 28 Feb 2024.
- [2] B-SCST: Bayesian Self-Critical Sequence Training for Image Captioning Shashank Bujimalla\* Intel Corporation shashankbvs@gmail.com Mahesh Subedar\* Intel Labs mahesh.subedar@intel.com Omesh Tickoo Intel Labs [omesh.tickoo@intel.com](mailto:omesh.tickoo@intel.com) 28 Jun 2020
- [3] SELF-CRITICAL SEQUENCE TRAINING FOR AUTOMATIC SPEECH RECOGNITION Chen Chen, Yuchen Hu, Nana Hou, Xiaofeng Qi, Heqing Zou, Eng Siong Chng 13 Apr 2022
- [4] Self-critical Sequence Training for Image Captioning Steven J. Rennie<sup>1</sup>, Etienne Marcheret<sup>1</sup>, Youssef Mroueh, Jerret Ross and Vaibhava Goel<sup>1</sup> Watson Multimodal Algorithms and Engines Group IBM T.J. Watson Research Center, NY, USA 16 Nov 2017
- [5] Explainable Semantic Communication for Text Tasks Chuanhong Liu, Caili Guo, Senior Member, IEEE, Yang Yang, Wanli Ni, Yanquan Zhou, Lei Li, and Tony Q.S. Quek, Fellow, IEEE 17 May 2024
- [6] A scoping review of transfer learning research on medical image analysis using ImageNet Mohammad Amin Morid, PhD<sup>1</sup>, Alireza Borjali, PhD<sup>2,3</sup>, Guilherme Del Fiol, MD, PhD<sup>4</sup> 13 Nov 2020
- [7] UNSUPERVISED REPRESENTATION LEARNING WITH DEEP CONVOLUTIONAL GENERATIVE ADVERSARIAL NETWORKS Alec Radford & Luke Metz indico Research Boston, MA {alec,luke}@indico.io Soumith Chintala Facebook AI Research New York, NY [soumith@fb.com](mailto:soumith@fb.com) 7 Jan 2016
- [8] Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks Jun-Yan Zhu\* Taesung Park\* Phillip Isola Alexei A. Efros Berkeley AI Research (BAIR) laboratory, UC Berkeley 24 Aug 2020
- [9] Generative Adversarial Nets Ian J. Goodfellow, Jean Pouget-Abadie\*, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair†, Aaron Courville, Yoshua Bengio‡ Departement d'informatique et de recherche opérationnelle † Université de Montréal ‡ Montréal, QC H3C 3J7 10 Jun 2014