

**THUYẾT MINH ĐỀ TÀI
NGHIÊN CỨU KHOA HỌC CẤP CƠ SỞ
(do sinh viên thực hiện)**

1. TÊN ĐỀ TÀI Xây dựng ứng dụng sinh câu mô tả cho hình ảnh trên nền tảng website và mobile.	2. MÃ SỐ THS2021-... (Phòng QLKH sẽ cấp khi đề tài được duyệt)						
Lĩnh vực ưu tiên <input type="checkbox"/> Lĩnh vực 1. Ứng dụng công nghệ cao trong nông nghiệp, thủy sản và môi trường <input type="checkbox"/> Lĩnh vực 2. Quản lý và sử dụng bền vững tài nguyên thiên nhiên <input checked="" type="checkbox"/> Lĩnh vực 3. Kỹ thuật công nghệ và công nghệ thông tin – truyền thông <input type="checkbox"/> Lĩnh vực 4. Khoa học Giáo dục, Luật và Xã hội Nhân văn <input type="checkbox"/> Lĩnh vực 5. Phát triển kinh tế, thị trường <input type="checkbox"/> Không thuộc 05 Lĩnh vực ưu tiên.							
3. LĨNH VỰC NGHIÊN CỨU Khoa học Tự nhiên <input type="checkbox"/> Khoa học Kỹ thuật và Công nghệ <input checked="" type="checkbox"/> X Khoa học Y, dược <input type="checkbox"/> Khoa học Nông nghiệp <input type="checkbox"/> Khoa học Xã hội <input type="checkbox"/> Khoa học Nhân văn <input type="checkbox"/>	4. LOẠI HÌNH NGHIÊN CỨU <table border="1" style="width: 100%; border-collapse: collapse; text-align: center;"><tr><td style="width: 33%;">Cơ bản</td><td style="width: 33%;">Ứng dụng</td><td style="width: 33%;">Triển khai</td></tr><tr><td> </td><td>X</td><td> </td></tr></table>	Cơ bản	Ứng dụng	Triển khai		X	
Cơ bản	Ứng dụng	Triển khai					
	X						
5. THỜI GIAN THỰC HIỆN 6 tháng Tù tháng 05 năm 2022 đến tháng 11 năm 2022							

6. ĐƠN VỊ CỦA CHỦ NHIỆM ĐỀ TÀI

Tên đơn vị: Khoa Công nghệ Thông tin và Truyền thông – Trường Đại học Cần Thơ

Điện thoại: 0292 3 734713 - 0292 3 831301

E-mail: office@cit.ctu.edu.vn

Địa chỉ: Khu 2, đường 3/2, Phường Xuân Khánh, Q. Ninh Kiều, TP. Cần Thơ, Việt Nam.

Họ và tên thủ trưởng đơn vị: Nguyễn Hữu Hòa

7. CHỦ NHIỆM ĐỀ TÀI

Họ và tên: Lê Anh Khôi

MSSV: B1910659

Ngày tháng năm sinh: 26/11/2001

Lớp: Công nghệ Thông tin - Chất lượng cao

Điện thoại di động: 0398635898

Khóa: 45

E-mail: khoib1910659@student.ctu.edu.vn

8. NHỮNG THÀNH VIÊN THAM GIA NGHIÊN CỨU ĐỀ TÀI

TT	Họ và tên	MSSV, Lớp, Khóa	Nội dung nghiên cứu cụ thể được giao	Chữ ký
1	Lê Anh Khôi	B1910659, Công nghệ thông tin chất lượng cao 01, khóa 45	<ul style="list-style-type: none">- Nghiên cứu lý thuyết về máy học, các mô hình neuron nhân tạo, nghiên cứu mô hình Inception V3 và mô hình LSTM- Nghiên cứu các bài báo liên quan đến sinh câu mô tả cho hình ảnh.- Lập trình, huấn luyện và đánh giá mô hình Inception V3 kết hợp với mô hình LSTM.	
2	Võ Triệu Vỹ	B1910730, Công nghệ thông tin chất lượng cao 01, khóa 45	<ul style="list-style-type: none">- Nghiên cứu mô hình InceptionV3- Nghiên cứu tập dữ liệu dùng trong việc huấn luyện mô hình.- Viết báo cáo tổng kết.	
3	Lê Đoàn Nhật Minh	B1900349, Công nghệ thông tin chất lượng cao 01, khóa 45	<ul style="list-style-type: none">- Nghiên cứu các bài báo liên quan đến sinh câu mô tả cho hình ảnh.- Nghiên cứu mô hình LSTM- Viết báo cáo tổng kết.	

4	Nguyễn Duy Khang	B1910652, Công nghệ thông tin chất lượng cao 01, khóa 45	<ul style="list-style-type: none"> - Nghiên cứu các bài báo liên quan đến sinh câu mô tả cho hình ảnh. - Nghiên cứu phương pháp đánh giá mô hình sinh câu mô tả cho hình ảnh - Viết báo cáo tổng kết. 	
5	Trần Bá Phương	B1908407, Công nghệ thông tin chất lượng cao 01, khóa 45	<ul style="list-style-type: none"> - Nghiên cứu lý thuyết về Inception V3 và LSTM. - Lập trình, huấn luyện và đánh giá mô hình InceptionV3 kết hợp với LSTM. -Nghiên cứu xây dựng website và ứng dụng mobile tích hợp mô hình sinh câu mô tả cho hình ảnh - Nghiên cứu tích hợp thư viện để đọc câu mô tả của ảnh thành giọng nói. 	

Cán bộ hướng dẫn sinh viên thực hiện đề tài

Họ và tên, MSCB	Đơn vị công tác và lĩnh vực chuyên môn	Nhiệm vụ	Chữ ký
TS. Lâm Nhựt Khang, MSCB: 1943	Đơn vị công tác: Khoa Công nghệ Thông tin và Truyền thông Lĩnh vực chuyên môn: Khoa học máy tính	Hướng dẫn nội dung khoa học và hướng dẫn lập dự toán kinh phí đề tài	

9. ĐƠN VỊ PHỐI HỢP CHÍNH

Tên đơn vị trong và ngoài nước	Nội dung phối hợp nghiên cứu	Họ và tên người đại diện đơn vị
Không	Không	Không

10. TỔNG QUAN TÌNH HÌNH NGHIÊN CỨU THUỘC LĨNH VỰC CỦA ĐỀ TÀI Ở TRONG VÀ NGOÀI NƯỚC

10.1. Trong nước

Trong những năm gần đây, nước ta đã thực hiện nhiều nghiên cứu liên quan đến lĩnh vực máy học. Các nghiên cứu về phương pháp xử lý hình ảnh, thị giác máy tính được sự quan tâm của nhiều tác giả. Về bài toán sinh câu mô tả cho hình ảnh cũng được thu hút sự chú ý của các diễn đàn chia sẻ kiến thức về công nghệ thông tin tại Việt Nam như Viblo.asia, tek4.vn,... và nhiều diễn đàn khác trong nước. Việc phát triển một sản phẩm ứng dụng các mô hình sinh câu mô tả cho hình ảnh vào thực tiễn chưa được nhiều người tiếp cận.

10.2. Ngoài nước

Nghiên cứu sinh câu mô tả cho hình ảnh là một bài toán rất được quan tâm ở các tạp chí công nghệ, bài báo nghiên cứu của các trường Đại học ngoài nước. Bài toán sinh câu mô tả thường bao gồm 2 giai đoạn là rút trích đặc trưng ảnh bằng mô hình dựa vào kỹ thuật thị giác máy tính và sinh câu mô tả cho hình ảnh bằng mô hình ngôn ngữ. Các mô hình dùng để rút trích đặc trưng ảnh như VGGNet [1], Faster R-CNN [2], ResNet [3], and Inception-V3 [26]. Mô hình ngôn ngữ có thể dụng các mô hình học sâu như LSTM [5,6], BiLSTM [7], và Transformer [8,9].

Tài liệu tham khảo:

- [1] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. arXivpreprint arXiv:1409.1556. 2014.
- [2] Ren S, He K, Girshick R, Sun J. Faster R-CNN: Towards real-time object detection with region proposal networks. Advances in Neural Information Processing Systems. 2015; 28. p.91-99.
- [3] He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition; 2016. p. 770-778.
- [4] Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z. Rethinking the inception architecture for computer vision. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; 2016. p.2818-2826.
- [5] YAO, Ting, et al. Boosting image captioning with attributes. In: Proceedings of the IEEE international conference on computer vision. 2017. p. 4894-4902.
- [6] Soh, Moses. "Learning CNN-LSTM architectures for image caption generation." Dept. Comput. Sci., Stanford Univ., Stanford, CA, USA, Tech. Rep (2016).
- [7] Lu, Huimin, Rui Yang, Zhenrong Deng, Yonglin Zhang, Guangwei Gao, and Rushi Lan. "Chinese image captioning via fuzzy attention-based DenseNet-BiLSTM." ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM) 17, no. 1s (2021): 1-18.
- [8] LIU, Wei, et al. Cptr: Full transformer network for image captioning. arXiv preprint arXiv:2101.10804, 2021.
- [9] MISHRA, Santosh Kumar, et al. Image captioning in Hindi language using transformer networks. Computers & Electrical Engineering, 2021, 92: 107114.

10.3. Danh mục các công trình đã công bố thuộc lĩnh vực của đề tài của chủ nhiệm và những thành viên tham gia nghiên cứu

a) Của chủ nhiệm đề tài: Không

b) Của các thành viên tham gia nghiên cứu: Không

11. TÍNH CẤP THIẾT CỦA ĐỀ TÀI

Chúng tôi hướng đến việc phát triển một ứng dụng website có thể phục vụ nhu cầu tìm kiếm những hình ảnh dựa trên một số đặc trưng cụ thể trong tập số lượng lớn các hình ảnh được người dùng tải lên bằng cách nhập các từ khóa liên quan đến bức ảnh đó. Việc đó được thực hiện bằng cách sinh ra câu mô tả cho tất cả hình ảnh mà người dùng tải lên, sau đó so khớp từ khóa tìm kiếm của người dùng với câu mô tả để lọc ra những hình ảnh phù hợp.

Ứng dụng web cũng đóng vai trò là một công cụ cho phép người dùng đăng tải các bức ảnh và trả lại cho người dùng câu mô tả về hình ảnh đó đính kèm file audio đọc thành tiếng để người dùng tải về phục vụ cho mục đích cá nhân.

Đối với nền tảng di động, những người gặp vấn đề về thị lực có thể dùng ứng dụng để chụp lại môi trường xung quanh và nghe được câu mô tả về bức ảnh đó.

12. MỤC TIÊU ĐỀ TÀI

Xây dựng ứng dụng sinh câu mô tả cho ảnh (ứng dụng website và ứng dụng di động) có giá trị thực tiễn.

13. ĐỐI TƯỢNG, PHẠM VI NGHIÊN CỨU

13.1. Đối tượng nghiên cứu

- Các hình ảnh, các mô hình máy học để sinh câu mô tả cho hình ảnh và phương pháp tích hợp mô hình vào website, ứng dụng di động.
- Nghiên cứu ứng dụng các thư viện để đọc câu mô tả thành giọng nói.
- Quy trình phát triển website và ứng dụng di động.

13.2. Phạm vi nghiên cứu

- Phạm vi nghiên cứu là mô hình máy học sinh câu mô tả ở ngôn ngữ tiếng Anh cho hình ảnh

14. CÁCH TIẾP CẬN, PHƯƠNG PHÁP NGHIÊN CỨU

14.1. Cách tiếp cận

- Nghiên cứu lý thuyết - xây dựng mô hình, thực nghiệm và đánh giá kết quả, xây dựng ứng dụng có tích hợp mô hình.

14.2. Phương pháp nghiên cứu

- Nghiên cứu tài liệu tham khảo các nghiên cứu liên quan ở trong nước và ngoài nước.
- Sử dụng Google Colab để xây dựng mô hình.

15. NỘI DUNG NGHIÊN CỨU VÀ TIẾN ĐỘ THỰC HIỆN

15.1. Nội dung nghiên cứu

Nội dung nghiên cứu được chia làm 5 giai đoạn chính bao gồm:

- 1) Nghiên cứu lý thuyết: Nội dung nghiên cứu lý thuyết tập trung vào tìm hiểu các tài liệu sau:
 - Kiến thức cơ sở về máy học, mạng nơ-ron
 - Nghiên cứu về mô hình Inception-V3, LSTM, cách kết hợp mô hình Inception-V3 với LSTM
 - Nghiên cứu phương pháp đánh giá mô hình sinh câu mô tả cho hình ảnh
 - Nghiên cứu các phương pháp và quy trình phát triển ứng dụng website và di động
- 2) Nghiên cứu dữ liệu huấn luyện:
 - Nghiên cứu các tập dữ liệu huấn luyện cho bài toán sinh câu mô tả có sẵn.
 - Tìm hiểu tập dữ liệu Flickr8k.
- 3) Xây dựng mô hình sinh câu mô tả cho hình ảnh.
- 4) Thực nghiệm và đánh giá kết quả.
- 5) Xây dựng ứng dụng website và ứng dụng di động.

15.2. Tiến độ thực hiện

STT	Các nội dung, công việc thực hiện	Sản phẩm	Thời gian (bắt đầu-kết thúc)	Người thực hiện và số ngày thực hiện
1.	Nghiên cứu lý thuyết về máy học; các nghiên cứu liên quan đến sinh câu mô tả cho hình ảnh và nghiên cứu phương pháp đánh giá mô hình sinh câu mô tả cho hình ảnh	Khoảng 5 - 7 trang A4 trình bày báo cáo lý thuyết về máy học, các nghiên cứu liên quan đến sinh câu mô tả cho hình ảnh ở trong và ngoài nước.	01/05/2022-01/06/2022	Trần Bá Phương (6 ngày); Lê Anh Khôi (9 ngày); Lê Đoàn Nhật Minh (9 ngày); Võ Triệu Vỹ (9 ngày); Nguyễn Duy Khang (9 ngày);
2.	Nghiên cứu dữ liệu huấn luyện.	Khoảng 2-3 trang A4 trình bày về các tập dữ liệu dùng cho bài toán sinh câu mô tả và tập dữ liệu Flickr8k.	01/6/2022	Trần Bá Phương (4 ngày)
3.	Xây dựng mô hình.	Khoảng 5-7 trang A4 trình bày về mô hình đã xây dựng.	01/06/2022-01/07/2022	Trần Bá Phương (10 ngày); Lê Anh Khôi (10 ngày); Võ Triệu Vỹ (10 ngày);
4.	Thực nghiệm và đánh giá kết quả	Khoảng 3-5 trang A4 trình bày cách triển khai thực nghiệm, kết quả	01/07/2022-08/07/2022	Võ Triệu Vỹ (6 ngày);

		thực nghiệm của mô hình.		Lê Đoàn Nhật Minh (6 ngày); Nguyễn Duy Khang (6 ngày);
5.	Xây dựng ứng dụng website và ứng dụng di động. Nghiên cứu tích hợp thư viện để đọc câu mô tả của ảnh thành giọng nói.	Khoảng 5-7 trang A4 trình bày về cách xây dựng website và ứng dụng di động áp dụng mô hình sinh câu mô tả cho hình ảnh.	08/07/2022- 30/09/2022	
6	Viết báo cáo tổng kết	Quyển báo cáo tổng kết	01/10/2022 30/10/2022	Võ Triệu Vỹ (10 ngày); Lê Đoàn Nhật Minh (10 ngày); Nguyễn Duy Khang (10 ngày);

16. SẢN PHẨM

Stt	Tên sản phẩm	Số lượng	Yêu cầu chất lượng sản phẩm (mô tả chi tiết chất lượng sản phẩm đạt được như nội dung, hình thức, các chỉ tiêu, thông số kỹ thuật,...)
I	Sản phẩm khoa học (Các công trình khoa học sẽ được công bố: sách, bài báo khoa học...): Không		
II	Sản phẩm đào tạo (Luận văn tốt nghiệp đại học): Không		
III	Sản phẩm ứng dụng: Ứng dụng sinh câu mô tả cho hình ảnh trên nền tảng website và mobile.		

17. PHƯƠNG THỨC CHUYỂN GIAO KẾT QUẢ NGHIÊN CỨU VÀ ĐỊA CHỈ ỨNG DỤNG

17.1. Phương thức chuyển giao

Chuyển giao cho Khoa Công Nghệ Thông Tin và Truyền Thông, Trường Đại Học Cần Thơ

17.2. Địa chỉ ứng dụng

Khoa Công nghệ Thông tin và Truyền thông, trường Đại học Cần Thơ.

18. TÁC ĐỘNG VÀ LỢI ÍCH MANG LẠI CỦA KẾT QUẢ NGHIÊN CỨU

18.1. Đối với lĩnh vực giáo dục và đào tạo

Giáo viên, giảng viên có thể sử dụng ứng dụng như một công cụ giảng dạy, minh họa cho sinh viên.

18.2. Đối với lĩnh vực khoa học và công nghệ có liên quan

Kết quả nghiên cứu có thể được sử dụng, cải tiến để làm cơ sở cho các nghiên cứu khoa học và công nghệ khác trong tương lai.

18.3. Đối với phát triển kinh tế-xã hội

Ứng dụng di động đóng vai trò như một người trợ lý hỗ trợ cho người khiếm thị với khả năng chụp ảnh và đọc mô tả thành tiếng.

18.4. Đối với tổ chức chủ trì và các cơ sở ứng dụng kết quả nghiên cứu

Góp phần thu hút sự quan tâm của sinh viên đến hoạt động nghiên cứu khoa học, khơi nguồn cảm hứng cho các ý tưởng, sáng tạo mới.

Ứng dụng có thể được tiếp tục phát triển trong tương lai để mang lại nhiều giá trị hơn cho cộng đồng và xã hội.

19. KINH PHÍ THỰC HIỆN ĐỀ TÀI VÀ NGUỒN KINH PHÍ

Kinh phí thực hiện đề tài: 15.000.000 đồng.

Trong đó:

Kinh phí Trường cấp: 15.000.000 đồng.

Các nguồn khác: 0 đồng.

Đơn vị tính: đồng

Stt	Khoản chi, nội dung chi	Tổng kinh phí	Nguồn kinh phí	
			Kinh phí Trường cấp	Các nguồn khác
1	Chi mua vật tư, nguyên, nhiên, vật liệu	0	0	0
2	Chi tiền công lao động trực tiếp	12.275.000	12.275.000	0
3	Chi văn phòng, phẩm, thông tin liên lạc, in ấn	0	0	0
4	Chi họp hội đồng đánh giá, nghiệm thu	2.725.000	2.725.000	0
	Tổng cộng	15.000.000	15.000.000	0

Ngày ... tháng 04 năm 2022

KHOA CNTT & TT

CÁN BỘ HƯỚNG DẪN

CHỦ NHIỆM ĐỀ TÀI

**TL.HIỆU TRƯỞNG
TRƯỞNG PHÒNG QUẢN LÝ KHOA HỌC**