
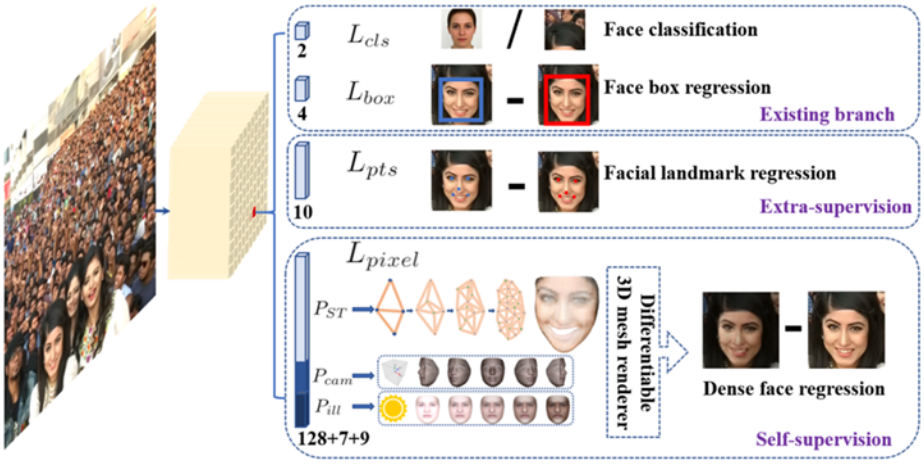
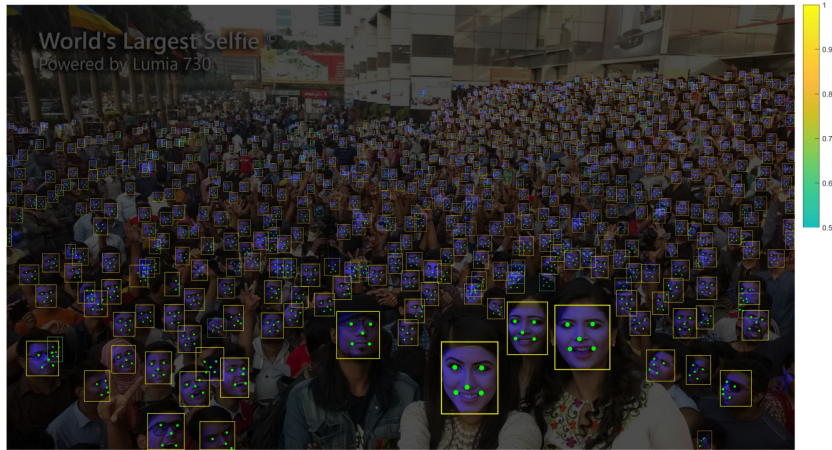


MẪU BÁO CÁO CỦA MỖI HV

Họ và tên (IN HOA)	NGUYỄN PHÚC THỊNH (CH2001038)
Ảnh	
Số buổi vắng	0
Bonus	5
Tên đề tài (VN)	Nhận diện nhiều khuôn mặt trong một bức ảnh tự nhiên
Tên đề tài (EN)	RetinaFace: Single-stage Dense Face Localisation in the Wild
Giới thiệu	<ul style="list-style-type: none">Ngày nay, trong một bức ảnh không chỉ là ảnh chân dung mà có khi là ảnh chụp của một nhóm từ vài người đến rất nhiều người. Trong một vài tình huống thực tế ta muốn nhận dạng từng người trong bức ảnh đó để chia sẻ những kỷ niệm hoặc là phục vụ trong công việc.Đề tài này trình bày một phương pháp single-stage phát hiện khuôn mặt một cách mạnh mẽ, được đặt tên là RetinaFace. RetinaFace thực hiện

	<p>định vị khuôn mặt thông minh theo pixel trên các quy mô khuôn mặt khác nhau.</p> <ul style="list-style-type: none"> Input của RetinaFace là những bức ảnh tự nhiên. Output là phân loại khuôn mặt, tạo ra Bounding box khuôn mặt, đánh dấu 5 điểm chính trên khuôn mặt (2 điểm ở 2 mắt, ngay giữa mũi và 2 khóe miệng), hình 3D và sự tương ứng của các pixel trên khuôn mặt.  <p>The diagram illustrates the RetinaFace architecture. It starts with an input image of a crowd, which is processed by a feature pyramid. The output branches are as follows:</p> <ul style="list-style-type: none"> Face classification: L_{cls} (2) with a loss of 2. It shows two face images with different classification results. Face box regression: L_{box} (4) with a loss of 4. It shows two face images with bounding boxes. Facial landmark regression: L_{pts} (10) with a loss of 10. It shows two face images with 5 landmark points each. Dense face regression: L_{pixel} (128+7+9) with a loss of 128+7+9. It shows a 3D mesh rendering process involving P_{ST}, P_{cam}, and P_{ill} components, leading to a differentiable 3D mesh renderer and dense face regression. <p>The existing branch includes Face classification and Face box regression. The extra-supervision branch includes Facial landmark regression. The self-supervision branch includes Dense face regression.</p>
<p>Mục tiêu</p>	<ul style="list-style-type: none"> Tác giả đánh dấu 5 điểm trên khuôn mặt của tập dữ liệu WIDER FACE và quan sát thấy sự cải thiện đáng kể trong nhận diện khuôn mặt. Tác giả sử dụng phương pháp tiên tiến nhất tập trung vào thiết kế một giai đoạn (single-stage), lấy mẫu dày đặc các vị trí và tỷ lệ của khuôn mặt dựa trên Feature Pyramid. Tác giả đã cải thiện khung nhận diện khuôn mặt một giai đoạn (single-stage) và sử dụng tính năng giảm thiểu tổn thất đa tác vụ, từ các tín hiệu được giám sát, để đề xuất phương pháp có thể phát hiện được khuôn mặt dù là rất nhỏ.



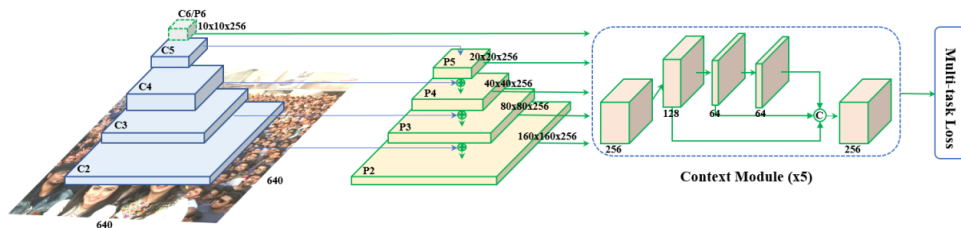
**Nội dung
và
phương
pháp thực
hiện**

- Nội dung nghiên cứu của bài báo này là đề cập đến một định nghĩa rộng hơn về bản địa hóa khuôn mặt, căn chỉnh khuôn mặt, phân tích cú pháp gương mặt theo pixel và hồi quy tương ứng dày đặc 3D. Loại định vị khuôn mặt dày đặc đó cung cấp thông tin chính xác về vị trí trên khuôn mặt cho tất cả các mô hình khác nhau.
- Thuật toán sử dụng cho Multi-task loss

$$L = L_{cls}(p_i, p_i^*) + \lambda_1 p_i^* L_{box}(t_i, t_i^*) + \lambda_2 p_i^* L_{pts}(l_i, l_i^*) + \lambda_3 p_i^* L_{pixel}.$$

Trong đó:

- P_i là dự đoán xác suất của I
- P_i^* là 1 nếu có khuôn mặt và 0 nếu không có khuôn mặt



Hình: Tổng quát về phương pháp bản địa hóa mật độ điểm trên mặt với single-stage. Sau các mô-đun ngữ cảnh, tác giả tính toán tổn thất đa tác vụ cho mỗi anchor

- Thuật toán Dense Regression loss.

$$L_{pixel} = \frac{1}{W * H} \sum_i^W \sum_j^H \|\mathcal{R}(\mathcal{D}_{P_{ST}}, P_{cam}, P_{ill})_{i,j} - I_{i,j}^*\|_1$$

Trong đó:

- W và H là chiều rộng và cao của bounding box
- $P_{cam} = [xc; yc; zc; x'c; y'c; z'c; fc]$
- $P_{ill} = [xl; yl; zl; rl; gl; bl; ra; ga; ba]$
- Tác giả sử dụng bộ dữ liệu WIDER FACE gồm có 32,203 bức ảnh và 393,703 bounding box khuôn mặt. Bộ dữ liệu được chia thành training (40%), validation (10%) và test (50%) bằng cách lấy ngẫu nhiên.
- Thực nghiệm với thí nghiệm cắt bỏ (Ablation) để kiểm tra xem 5 mốc điểm trên khuôn mặt được chú thích (five facial landmarks) và nhánh mật độ hồi quy (Dense Regression Branch) được đề xuất ảnh hưởng định lượng như thế nào đến hiệu suất nhận diện khuôn mặt.
- Thực nghiệm với thí nghiệm xác định bounding box cho khuôn mặt.
- Thực nghiệm đo độ chính xác của 5 điểm trên khuôn mặt (Five facial landmark)
- Thực nghiệm đo độ chính xác của mật độ điểm trên khuôn mặt. Với thí nghiệm này, RetinaFace xuất ra tương ứng khuôn mặt dày đặc, nhưng nhánh hồi quy dày đặc chỉ được đào tạo bằng cách học tự giám sát.

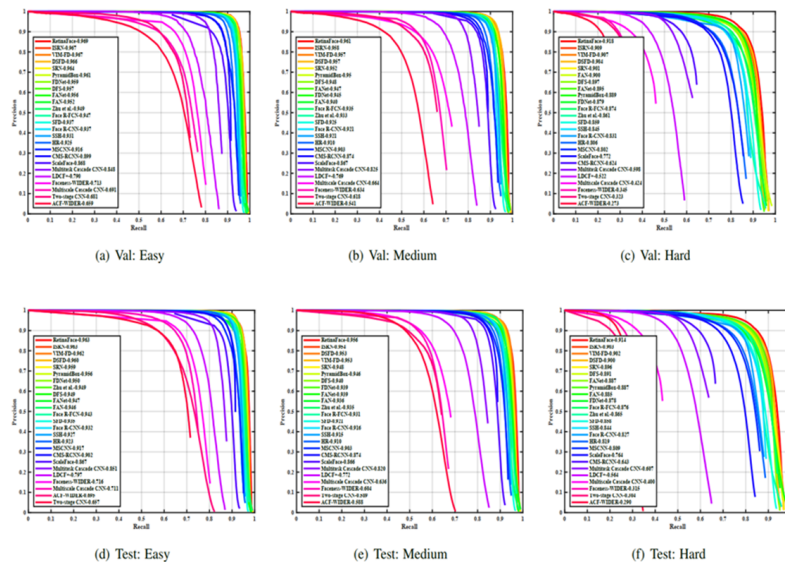
**Kết quả
dự kiến**

- Kết quả của thí nghiệm cắt bỏ (Ablation)

Method	Easy	Medium	Hard	mAP [33]
FPN+Context	95.532	95.134	90.714	50.842
+DCN	96.349	95.833	91.286	51.522
+ L_{pts}	96.467	96.075	91.694	52.297
+ L_{pixel}	96.413	95.864	91.276	51.492
+ $L_{pts} + L_{pixel}$	96.942	96.175	91.857	52.318

Tác giả đánh giá performance của vài cái setting khác nhau trên bộ data WIDER FACE. Tác giả thiết lập baseline là 91,286%. Với + L_{pts} + L_{pixel} thì kết quả tốt nhất trên cả 3 subset: Easy, Medium và Hard và mean Average Precision.

- Kết quả độ chính xác của Face Box



Kết quả của Retina Face là tốt nhất trên tất cả các subet Easy, Medium và Hard với 96.9% (Easy), 96.1% (Medium) and 91.8% (Hard) cho bộ Validation set; 96.3% (Easy), 95.6% (Medium) and 91.4% (Hard) cho test set.

- Kết quả độ chính xác của 5 điểm trên khuôn mặt (Five facial landmark)

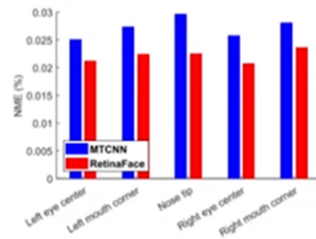


Fig. (a) NME on AFLW

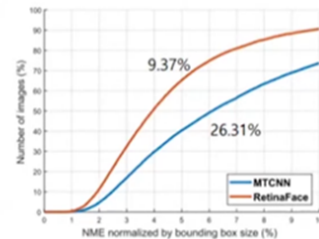
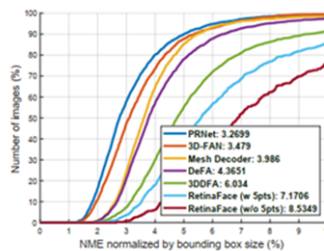


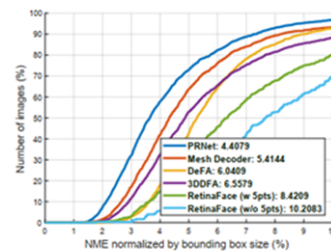
Fig. (b) CED on WIDER FACE

Để đánh giá độ chính xác của 5 điểm trên khuôn mặt, tác giả đã so sánh RetinaFace với MTCNN trên bộ dữ liệu AFLW (24,386 khuôn mặt) với bộ dữ liệu validation của WIDER FACE (18,5K khuôn mặt).

- Fig. (a) mean errors của MTCNN đều cao hơn RetinaFace ở tất cả vị trí của 5 điểm trên khuôn mặt.
- fig. (b) đường cong cumulative error distribution (CED) của MTCNN cao hơn đáng kể so với RetinaFace.
- Kết quả độ chính xác của mật độ điểm trên khuôn mặt (Dense Regression Branch).



(a) 68 2D Landmarks



(b) All 3D Landmarks

Hồi quy năm mốc trên khuôn mặt có thể giảm bớt khó khăn trong đào tạo của nhánh hồi quy dày đặc và cải thiện đáng kể kết quả hồi quy dày đặc.

Sử dụng các tính năng một giai đoạn (như trong RetinaFace) để dự đoán các thông số tương ứng dày đặc khó hơn nhiều so với việc sử dụng các tính năng RoI (Khu vực quan tâm).

- Kết quả độ chính xác của nhận diện khuôn mặt

Methods	LFW	CFP-FP	AgeDB-30
MTCNN + ArcFace	99.83	98.37	98.15
RetinaFace + ArcFace	99.86	99.49	98.60

Trong bảng, tác giả cho thấy ảnh hưởng của việc căn chỉnh và phát hiện khuôn mặt đối với nhận dạng khuôn mặt sâu (tức là ArcFace) bằng cách so sánh MTCNN được sử dụng rộng rãi và RetinaFace được đề xuất. Kết quả trên CFP-FP chứng minh rằng RetinaFace có thể tăng độ chính xác xác minh của ArcFace từ 98: 37% lên 99: 49%.

**Tài liệu
tham
khảo**

1. Ministry of Works (1962) Survey of Problems Before the Construction Industries, HMSO [online] available at: <https://archive.org/details/op1265587-1001> [Accessed 15 April 2020]
2. Banwell C.H., (1964) The Placing and Management of Contracts for Building and Civil Engineering Work, HMSO [online] available at: <https://archive.org/details/op1265594-1001> [Accessed 15 April 2020]
3. EDCs for Building and Civil Engineering (1975) The Public Client and the Construction Industries (The Wood Report), HMSO
4. Latham M, (1994) Constructing the Team, HMSO [online] available at: <http://constructingexcellence.org.uk/wpcontent/uploads/2014/10/Constructing-the-team-The-Latham-Report.pdf> [Accessed 15 April 2020]
5. Construction Task Force (1998) Rethinking Construction, HMSO [online] available at: http://constructingexcellence.org.uk/wp-content/uploads/2014/10/rethinking_construction_report.pdf [Accessed 15 April 2020]
6. Constructing Excellence (2009) Never Waste A Good Crisis, [online] available at: <https://constructingexcellence.org.uk/resources/never-waste-a-good-crisis/> [Accessed 15 April 2020]
7. Cabinet Office (2011) Government Construction Strategy, [online] available at: <https://www.gov.uk/government/publications/government-construction-strategy> [Accessed 15 April 2020]
8. British Standards Institution (2019) BS EN ISO 19650 Organization and digitization of information about buildings and civil engineering

	<p>works, including building information modelling - Information management using building information modelling, London: BSI</p> <p>9. Construction Industry Council (2018) Building Information Model (BIM) Protocol, 2nd Edition, London: CIC, [online] available at: http://cic.org.uk/admin/resources/bim-protocol-2nd-edition-2.pdf [Accessed 15 April 2020]</p> <p>10. Construction Industry Council (2013) Outline Scope of Services for the Role of Information Management, 1st Edition, London: CIC [online] available at: http://cic.org.uk/download.php?f=outline-scope-of-services-for-the-roleof-information-managment.pdf [Accessed 15 April 2020]</p> <p>11. British Standards Institution (2014) BS 1192-4:2014 Collaborative production of information. Fulfilling employer's information exchange requirements using COBie. Code of practice, London: BSI</p>
--	--