28th - International Conference on Pattern Recognition
Lyon, France August, 17-22, 2026

SPOT-Face: Forensic Face Identification using Attention Guided Optimal Transport

Ravi Shankar Prasad, Dinesh Singh
Visual Intelligence and Machine Learning Group (VIML)
School of Computing and Electrical Engineering (SCEE)
Indian Institute of Technology Mandi, India

Content

1. Introduction
2. Related Work
3. Dataset
4. Methodology
5. Experimental Results
6. Conclusion and Discussion

Introduction

Craniofacial Identification

- Craniofacial means skull and the face (Cranium→ Skull, Facial→ Face).
- 50,000 individuals remains unidentified every year[1]. (● Crime victims ● disaster victims)
- Identification of deceased individual using skull.

February 21, 1985 — FORENSIC MEDICINE — Hospital Doctor 13

Grim 'jigsaw' solved murders

Prof Bernard Knight looks at a classic example of how the expertise of medical professors brought a doctor to the gallows

ONE of the classic exercises in anatomical reconstruction was perhaps appropriately made in a double murder committed by a medical practitioner.

A woman out for a stroll one Sunday in 1955 was shocked to see a human arm sticking from a parcel thrown into a small stream just over the Scottish border, on the Carlisle-Edinburgh road.

After a police search, no less than 70 fragments of human remains were recovered, wrapped clumsily in newspaper and cloth.

Within the next two months, more pieces of limb were found scattered around the countryside.

The mutilated parts were taken to Glasgow, where they were examined by John Glaister, professor of forensic medicine, and J. C. Brash, professor of anatomy. They were later assisted by Professor Sydney Smith from Edinburgh.

They had to deal with two heads, two incomplete trunks, seven sections of arms and hands, four legs with only two feet and a large mass of flesh including three female breasts, genitalia and a uterus.

It was obvious that a deliberate effort had been made to render identification impossible, as the facial elements such as eyes, lips, noses and ears were missing and all teeth had been extracted after death, as had the finger-pads carrying prints.

The photograph of Mrs Ruxton used by detectives in the first recorded 'photo-fit' experiment

The photograph negative of Mrs Ruxton's skull which police fitted perfectly over her photograph

Painstaking work was done to assemble as much as possible, which it became evident that two female bodies were represented.

One was a young woman between 18 and 25, about 4ft 11in, with an appendix scar, eight teeth missing before death and chronic conilitis.

The other was an older woman, probably in her thirties, 5ft 3in, having borne children and suffered from a bunion.

While the doctors had been working, the police had also been active. Among the wrappings of the grisly parcels was a copy of the *Sunday Graphic* for September 15. This contained a report of the crowning of a carnival queen in Moreambe.

Detectives soon found this was a special local edition for the Lancaster area.

At about the same time the chief constable of Dumfries recalled reading a newspaper report that the housemaid of a Lancaster doctor had gone missing. When the Lancashire police made inquiries, they found that the doctor's wife was also missing.

The doctor in question was an Indian, Bakhtyar Rustomji Hakim, a Parsee family practitioner, who had qualified in

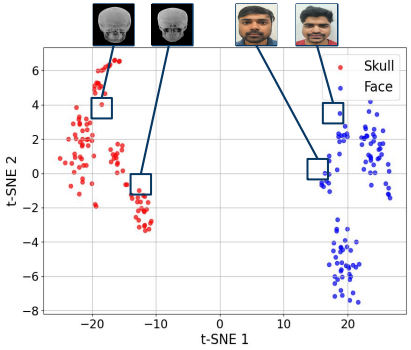
alamy — Image ID: GA7HME — www.alamy.com

[1] Behera, Chittaranjan, et al. "Development of the first DNA database and identification portal for identification of Unidentified bodies in India–UMID." *Science & Justice* 62.1 (2022): 110-116.

Problems and Objectives

Problems:

- ➔ Very high domain-gap.
- ➔ Skull don't have visible attribute like face.
- ➔ Limited availability of benchmark skull-to-face dataset.

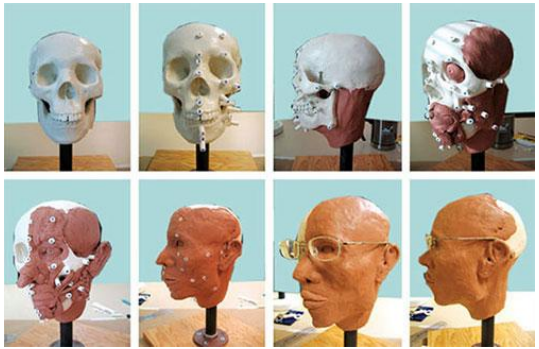


Objectives:

- To create craniofacial dataset.
- To develop novel method for craniofacial identification (either by matching or by reconstruction) from partial human remains in forensic investigations.

Existing Methods in Craniofacial Identification

Traditional way



Clay based method[2]

Limitations:

1. Based on tissue depth mark information.
2. Time consuming.
3. Need expert.

Superimposition based method[3]

Limitations:

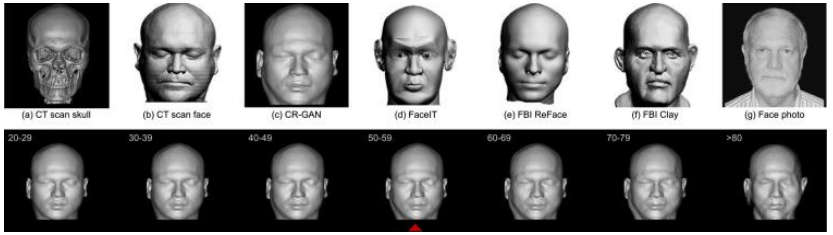
1. Based on tissue depth mark information.
2. Manual process of overlapping skull onto face.

[2] Taylor, Karen T. *Forensic art and illustration*. CRC press, 2000.
[3] Damas, S. et al. Forensic identification by computer-aided craniofacial superimposition: a survey. ACM Comput. Surv.(CSUR) 43, 1–27 (2011)

Continued..

Modern deep learning way

Direct skull to face matching method[4]



Direct skull to face generation method[5]

Limitations:

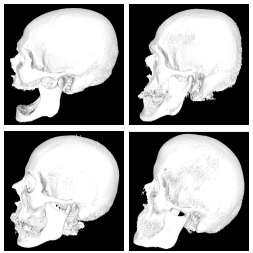
1. Data is collected from various sources.
2. HOG features are used for matching
3. Low accuracy

Limitations:

1. Data is collected from 3D CT scans.
2. Baised to chinese population.
3. Training instability.

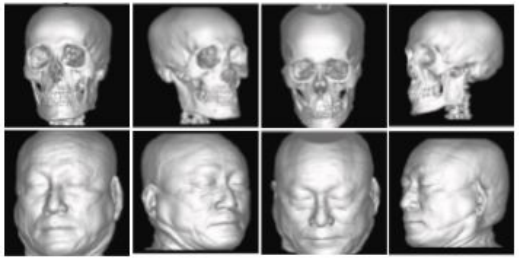
[4] Singh, Maneet, et al. "Learning a shared transform model for skull to digital face image matching." *2018 IEEE 9th International Conference on Biometrics Theory, Applications and Systems (BTAS)*. IEEE, 2018.
[5] Li, Yuan, et al. "CR-GAN: Automatic craniofacial reconstruction for personal identification." *Pattern Recognition* 124 (2022): 108400.

Craniofacial Datasets



MUG500+[6]

IdentifyMe[4]



Chinese CT dataset[5]

IITMandi_S2F (Ours)

Dataset	# Subjects	# Paired Samples	Skull Modality	Face Modality	Pose/View	Public
MUG500+	500	×	CT-derived Skull	×	Multi-view	✓
IdentifyMe	464	464	Skull Photograph	Face Photograph	Frontal	✓
ICCR-Diff	4802	72016	CT-derived Skull	CT-derived Face	Multi-view	×
IITMandi_S2F (Ours)	51	5100	Skull X-ray	RGB Face Photograph	Frontal + Lateral	✓

[6] Li, Jianning, et al. "MUG500+: Database of 500 high-resolution healthy human skulls and 29 craniotomy skulls and implants." *Data in Brief* 39 (2021): 107524.

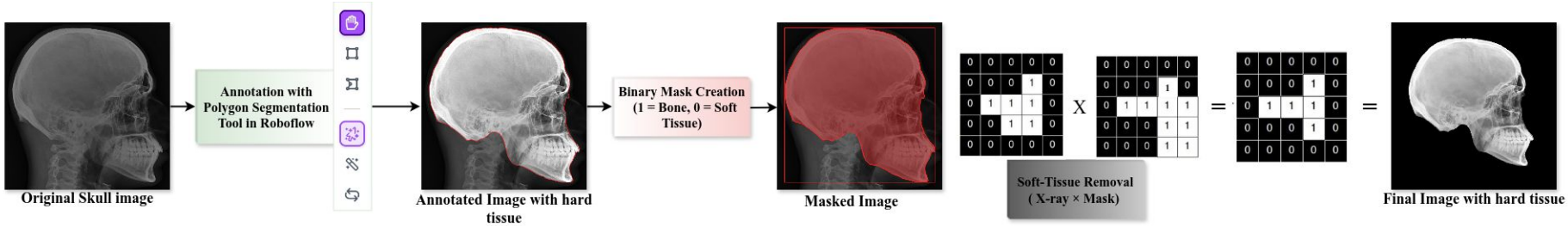
Methodology

Internationally recognized forensic artist at National center for Missing & Exploited Children, USA

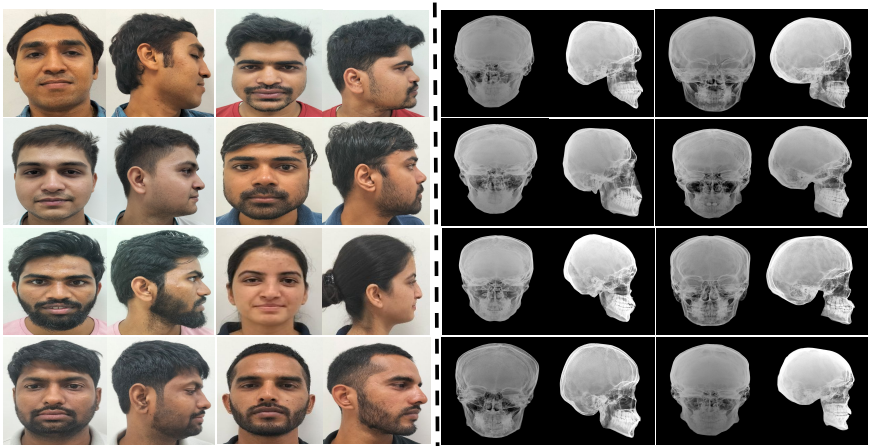
Quote

> "**Some say, 'That's such an old case, why do it?' I say, 'Why not? They still deserve their names back."**
> - Joe Mullins, National Center for Missing and Exploited Children.

IITMandi_S2F Datasets



Preprocessing step



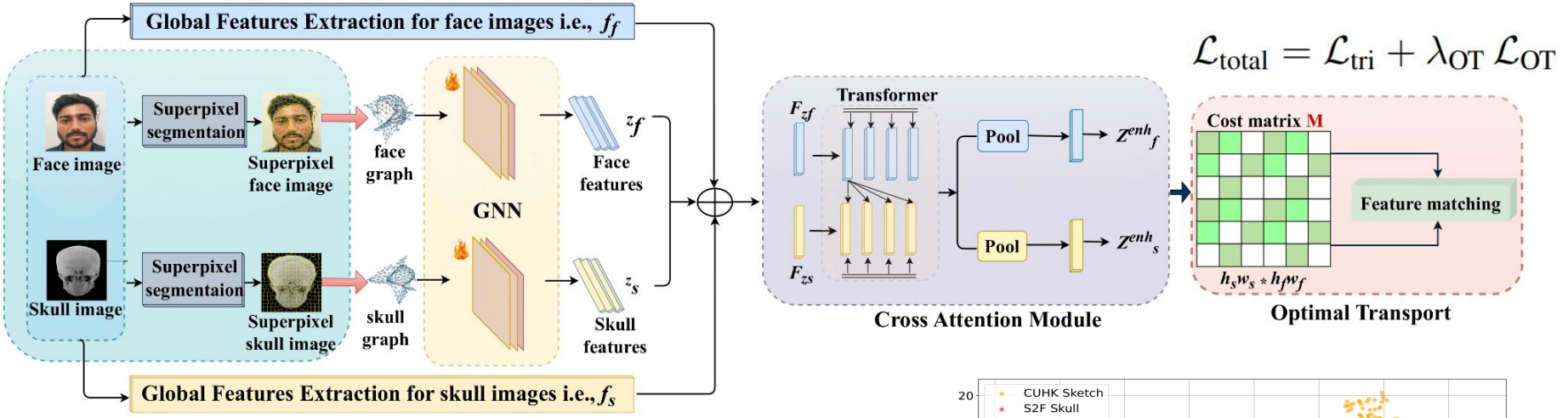
IITMandi_S2F

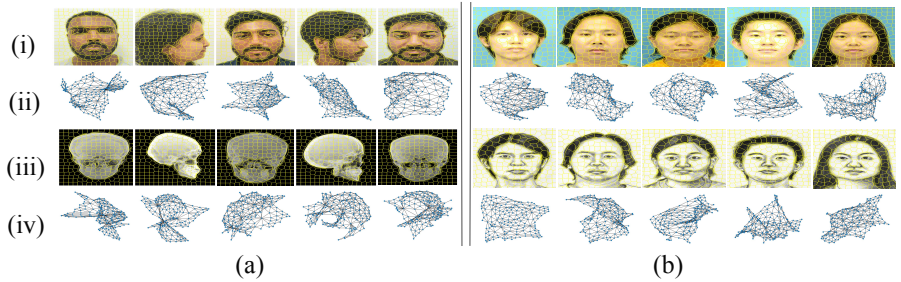
Contains X-ray image of face and their corresponding optical images from frontal and lateral views.

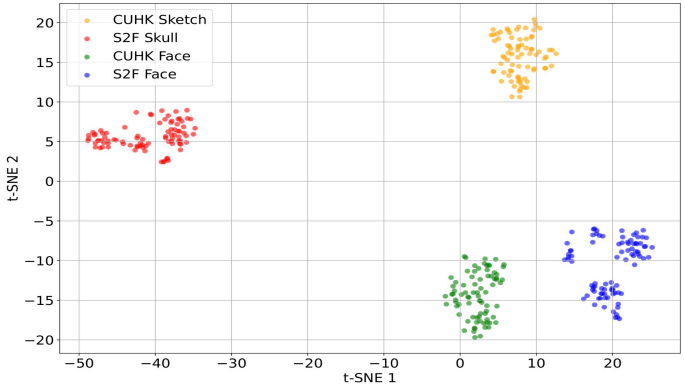
Total Subjects: 51 (Age: 21-30)
Male: 40
Female: 11

Craniofacial Identification with Matching

Title: *SPOT-Face: Forensic face identification using attention-guided optimal transport*[7]







[7] Prasad, R. S., & Singh, D. (2026). *SPOT-Face: Forensic face identification using attention-guided optimal transport*. (**Accepted in ICPR 2026**)

2: Quantitative Results[7]

Comparison of different GNN backbones on Skull–Face (IIT_Mandi_S2F) and Sketch–Face (CUFS) datasets. Here, Recall@K and mAP@K (%) are reported and the best results are shown in bold.

Skull–Face (IIT_Mandi_S2F)				
Backbone	**R@1**	**R@5**	**mAP@1**	**mAP@5**
GCN	50.0	71.0	50.0	63.9
GAT	49.5	68.1	49.5	62.0
GraphSAGE	50.0	70.1	50.0	63.4
GraphTransformer	**50.0**	**72.0**	**50.0**	**65.4**
Sketch–Face (CUFS)				
Backbone	**R@1**	**R@5**	**mAP@1**	**mAP@5**
GCN (Quickshift) [8]	82.4	–	–	–
GCN (SLIC) [8]	87.7	–	–	–
GCN	88.4	88.9	88.4	88.8
GAT	88.4	88.6	88.4	88.6
GraphSAGE	88.4	88.8	88.4	88.6
GraphTransformer	**88.4**	**89.0**	**88.4**	**88.9**

Impact of graph KNN hyperparameter (k) on Skull–Face (IIT_Mandi_S2F) and Sketch–Face (CUFS) datasets. Here, Recall@K and mAP@K (%) are reported and the best results are shown in bold.

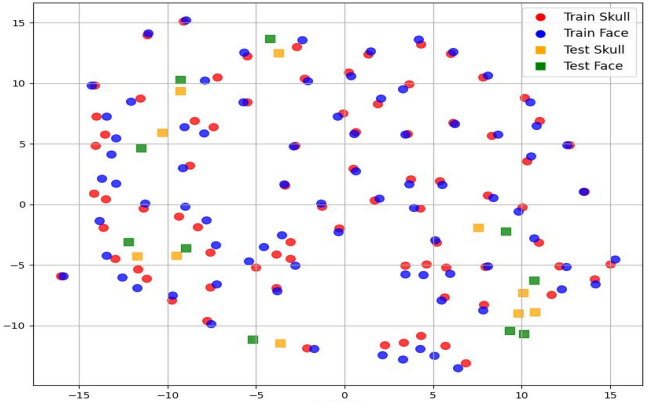
k	Skull–Face (S2F)				Sketch–Face (CUFS)			
	R@1	R@5	mAP@1	mAP@5	R@1	R@5	mAP@1	mAP@5
4	50.0	71.0	50.0	64.5	88.4	89.0	88.4	88.7
6	**50.0**	**72.0**	**50.0**	**65.4**	88.4	88.9	88.4	88.7
8	49.0	66.1	49.0	60.6	88.4	89.0	88.4	88.8
12	50.0	72.0	50.0	64.1	**88.4**	**89.0**	**88.4**	**88.9**

Ablation of Cross-Attention (CA) and Optimal-Transport (OT) modules on Skull–Face (S2F) and Sketch–Face (CUFS) matching. Recall@K and mAP@K (%) are reported. Best results are shown in bold.

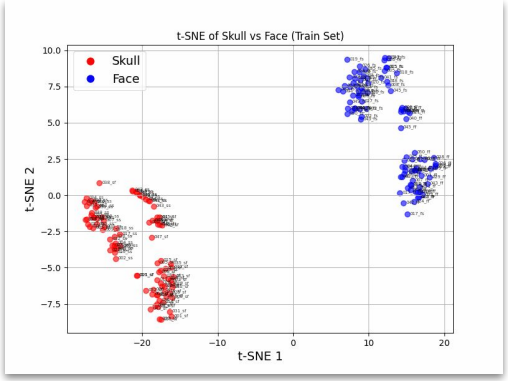
CA	OT	Skull–Face (S2F)				Sketch–Face (CUFS)			
		R@1	R@5	mAP@1	mAP@5	R@1	R@5	mAP@1	mAP@5
✗	✗	31.3	46.0	31.3	38.1	83.4	88.7	83.4	85.4
✓	✗	48.5	60.7	48.5	56.8	88.4	88.8	88.4	88.7
✗	✓	31.3	45.5	31.3	37.7	83.4	88.8	83.4	85.8
✓	✓	**50.0**	**72.0**	**50.0**	**65.4**	**88.4**	**89.0**	**88.4**	**88.9**

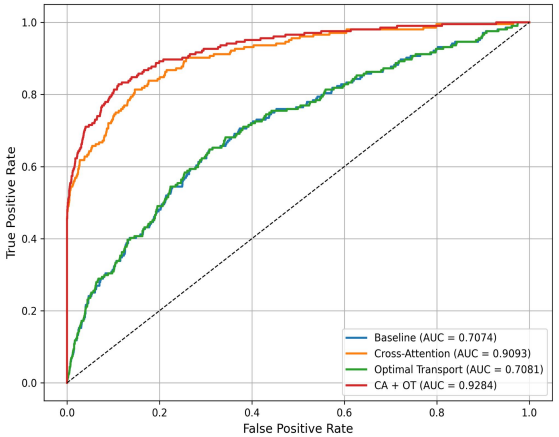
[8] Nagpal, Shruti, et al. "On matching skulls to digital face images: A preliminary approach." *2017 IEEE International Joint Conference on Biometrics (IJCB)*. IEEE, 2017.

2: Qualitative Results[7]

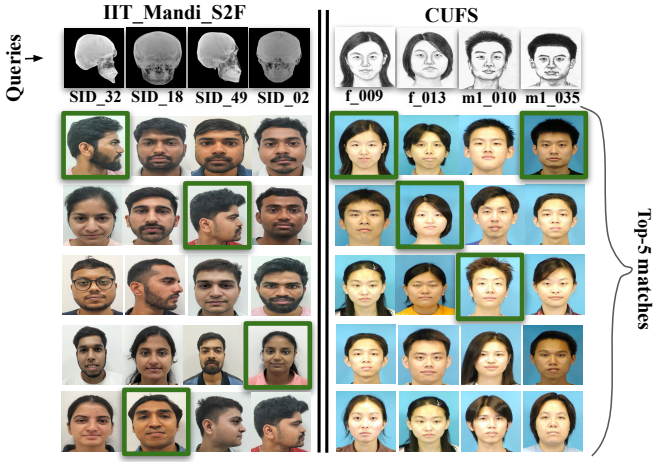


2D t-SNE visualization of skull and face embedding after training.





ROC-AUC curve



Top-5 retrieval results

Conclusion and Discussion

- We introduced a framework SPOT-Face for craniofacial identification with matching using 2D images of skulls and faces, including both frontal and lateral views.
- Our proposed method illustrates improved performance over previous methods, providing an alternative to manual annotation.
- Our results indicate that our proposed method performs significantly better in sketch-to-face matching, demonstrating its general applicability in various matching scenarios.
- Future work will explore how graph-based methods can be utilized for craniofacial reconstruction.

Thank You

Corresponding author

Co-author

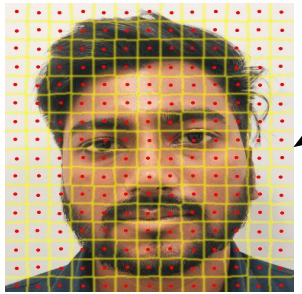
Google scholar

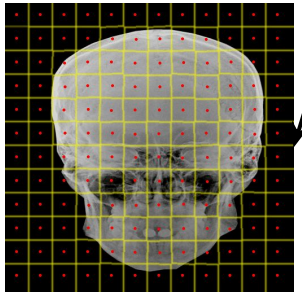
ResearchGate

LinkedIn

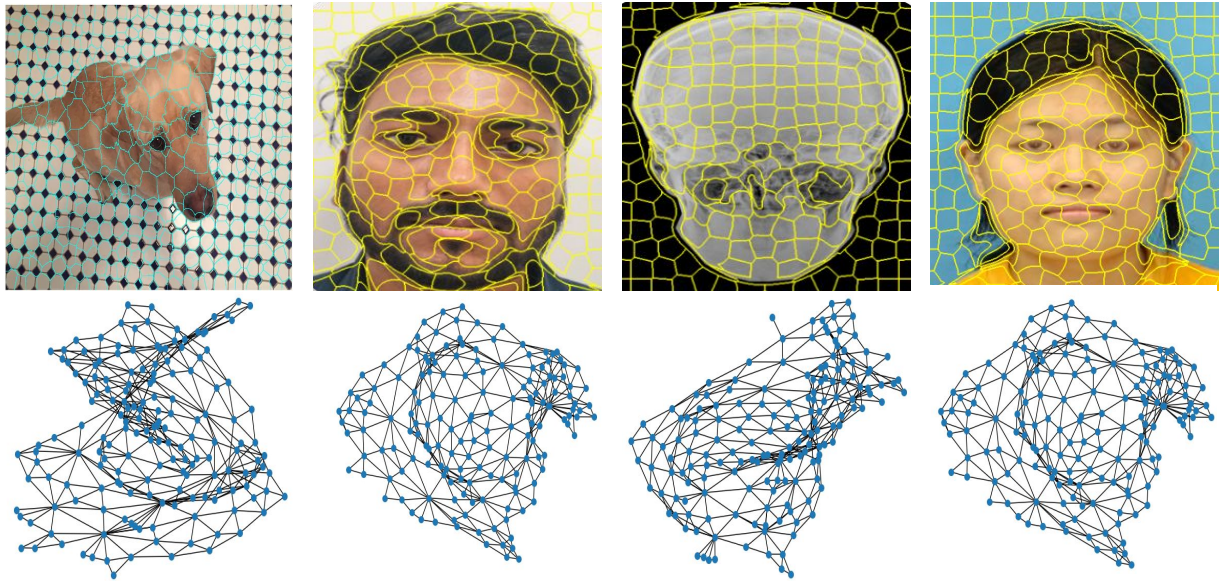
Supplementary slides

How to capture more structural information? SLIC[9] superpixel segmentation to Graph





SLIC: Simple Linear Iterative Clustering

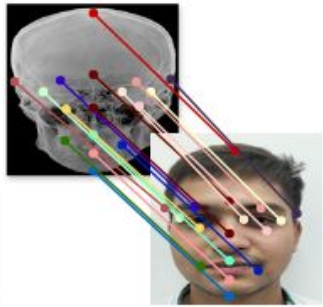


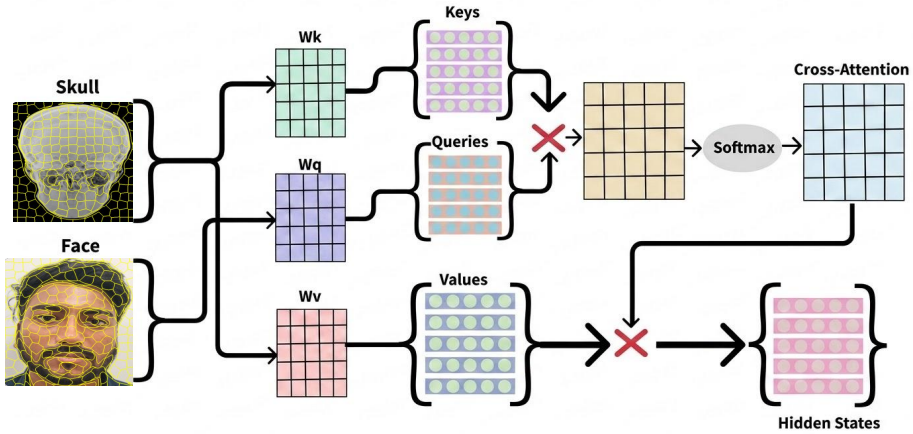
Node: each superpixel (R, G, B, x, y)

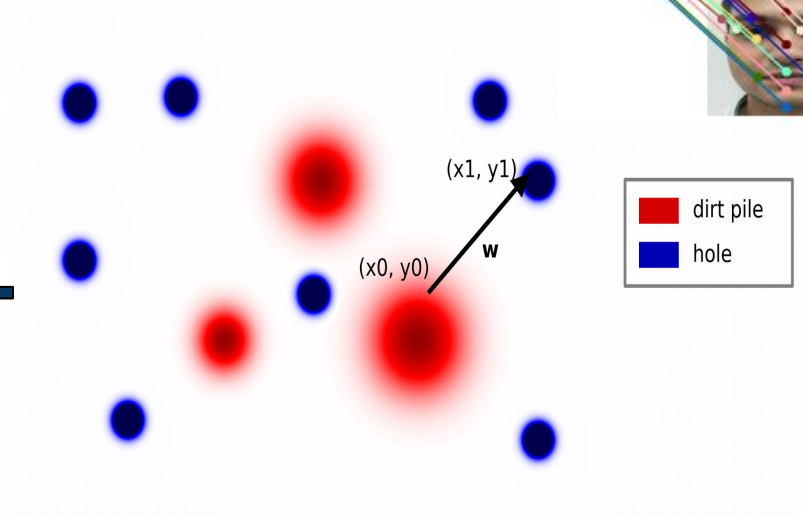
Edge: connection between each neighboring superpixel (KNN neighbour)

[9] Achanta, Radhakrishna, et al. "SLIC superpixels compared to state-of-the-art superpixel methods." *IEEE transactions on pattern analysis and machine intelligence* 34.11 (2012): 2274-2282.

How to get meaningful semantic correspondence? Cross-attention with Optimal Transport







Cross-Attention[10] (CA)

- ➔ Face has skin texture, eyes, nose, mouth.
- ➔ Skull has bone structure, eye socket, nasal bone.

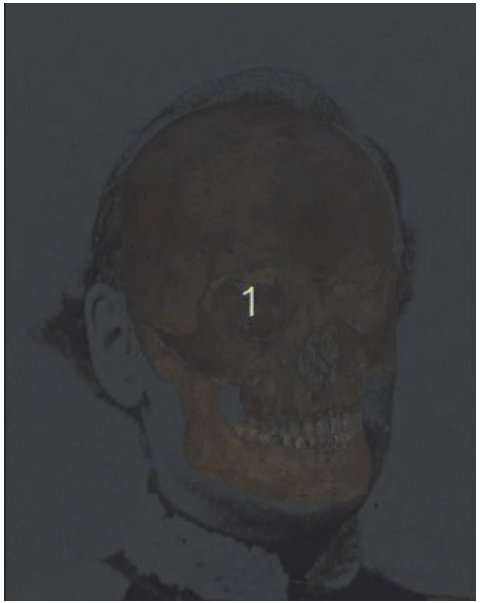
CA address→Which skull features should a face feature look at to become more informative?

Optimal Transport[11] (OT)

- ➔ Nasal bone → nose face = low cost.
- ➔ Eye socket → jaw face = high cost.

OT address→What is the best overall matching between skull regions and face regions?

[10] Vaswani, Ashish, et al. "Attention is all you need." *Advances in neural information processing systems* 30 (2017).
[11] Kantorovich, Leonid V. "On the Translocation of Masses." *Journal of mathematical sciences* 133.4 (2006).


1