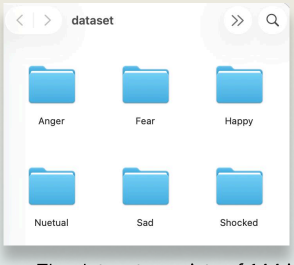
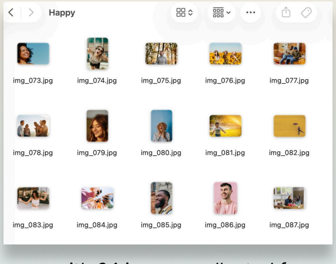
Facial Dataset

INTRODUCTION

The dataset contains **144 facial** images organized through a cross-structure of emotion.
It employs dual labeling (visual and contextual) along with metadata annotation to make the logic of data construction explicit.





The dataset consists of 144 images, with 24 images collected for each of the six emotion categories.

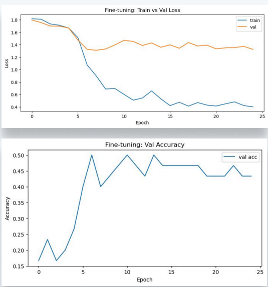
	A	B	C	D	E	F	G	H
1	**image_id**	**emotion**	**context**	**eyes**	**mouth**	**eyebrow**	**intensity**	**social_state**
2	img_001	sad	personal	downcast	open	furrowed	3	isolated
3	img_002	sad	personal	downcast	downturned	furrowed	2	isolated
4	img_003	sad	personal	closed	open	furrowed	3	isolated
5	img_004	sad	personal	downcast	downturned	furrowed	2	isolated
6	img_005	sad	personal	downcast	open	neutral	2	reactive
7	img_006	sad	personal	wide	neutral	neutral	2	reactive
8	img_007	sad	personal	downcast	downturned	neutral	2	isolated
9	img_008	sad	personal	downcast	neutral	neutral	1	reactive
10	img_009	sad	public	downcast	neutral	neutral	1	isolated
11	img_010	sad	personal	narrow	downturned	furrowed	2	isolated

Emotion labels were manually annotated for each image in an Excel metadata sheet.

APPROACH 1

Decontextualizing Emotion Computation

This approach uses a controlled comparison: it runs the ResNet18-based emotion model on original images and decontextualized images (cropped face, grayscale), then compares label, confidence, and accuracy changes to measure how context affects emotion recognition.



Mild overfitting occurs after early epochs


```
crop method: haar_face
pred_1 original: Anger (conf=0.6521)
pred_2 cropped face: Shocked (conf=0.3256)
pred_3 Grayscale face: Sad (conf=0.7378)
```

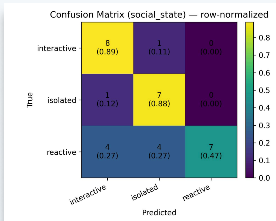
Predictions change under contextual removal

by Yirun Ye

APPROACH 2

Computing Social State from Facial Structure

This approach explores whether social behaviour can be computationally inferred from facial geometry, emotion labels, and contextual metadata, reframing emotion recognition as a model of social interpretation.



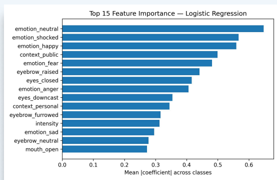
Confusion Matrix

Dataset + Model

127 valid samples
3 classes: interactive / isolated / reactive
Accuracy: 0.69

Findings

Interactive / isolated → stable prediction
Reactive → high confusion
Social state overlaps in facial cues
Not purely visual → context matters



Feature Importance

Key signals

emotion
context
eyebrow movement
eye closure

Social behaviour emerges from combined cues, not a single feature.

by Shuran Zhang

SHARED OBSERVATION

Both approaches reveal that facial interpretation is not purely visual.
When context is removed (Approach 1), emotion predictions shift and confidence decreases.
When context and metadata are structured (Approach 2), social_state becomes partially predictable but remains ambiguous in transitional cases.
This suggests that facial meaning is not fixed within the image itself.
Rather, it emerges from the interaction between **facial cues, contextual information, and computational modelling**.
Machine learning here does not simply recognise emotion — it operationalises how social interpretation is constructed.

ACKNOWLEDGEMENTS

Dataset construction

Yirun collected and labelled the Anger, Fear, and Sad image sets.
Shuran collected and labelled the Happy, Neutral, and Shocked image sets, and added the Social_State tags.

Poster Production

Yirun is responsible for developing the overall poster framework, including the introduction and individual section.
Shuran is responsible for organizing the observation findings and contributing her individual section.