Minimal Images Class

Shimon Ullman

Atoms of Recognition in Human and Computer Vision

Efficient use of limited information: recognizing local configurations

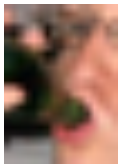




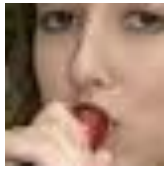
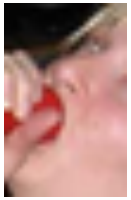



Minimizing variability





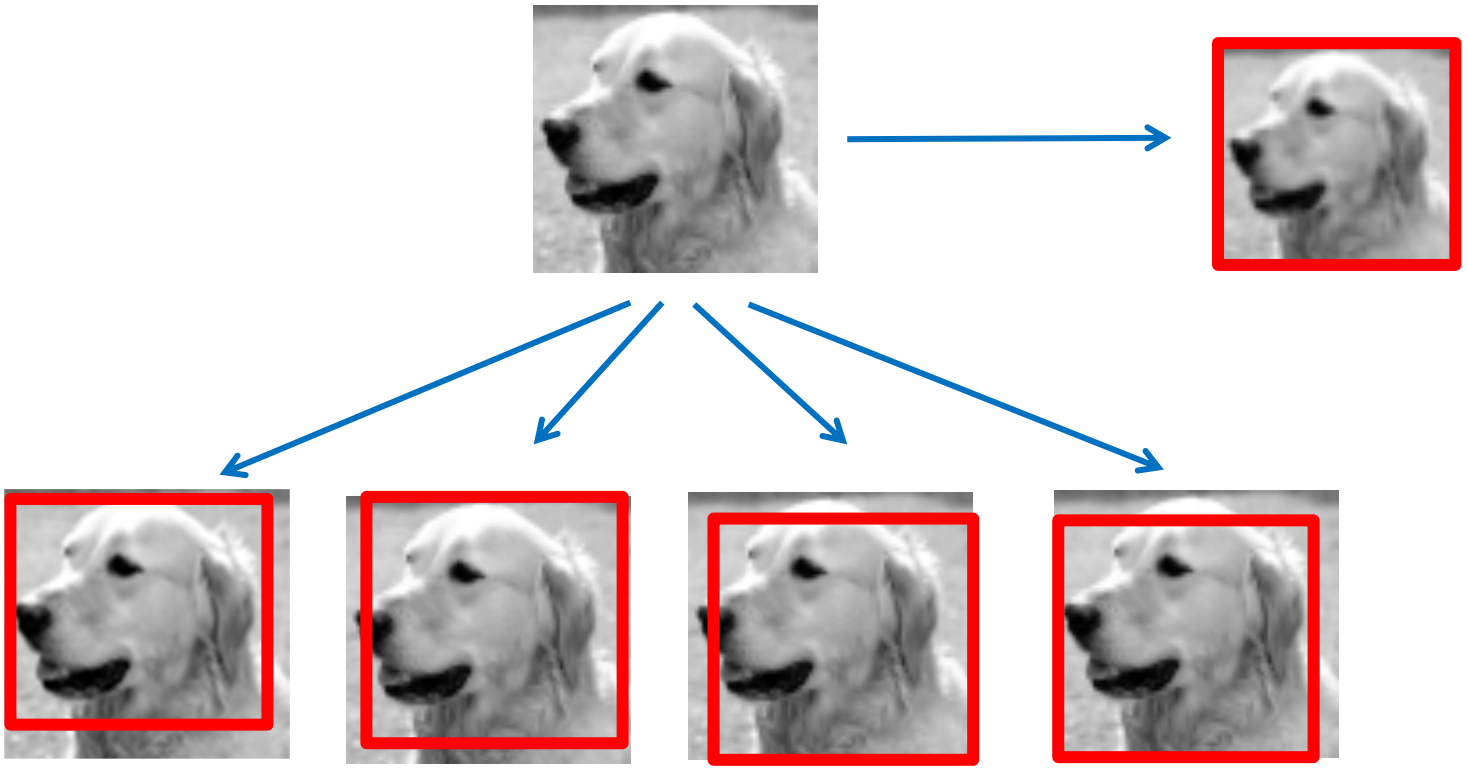



Minimal images provide a useful for the interpretation of complex scenes

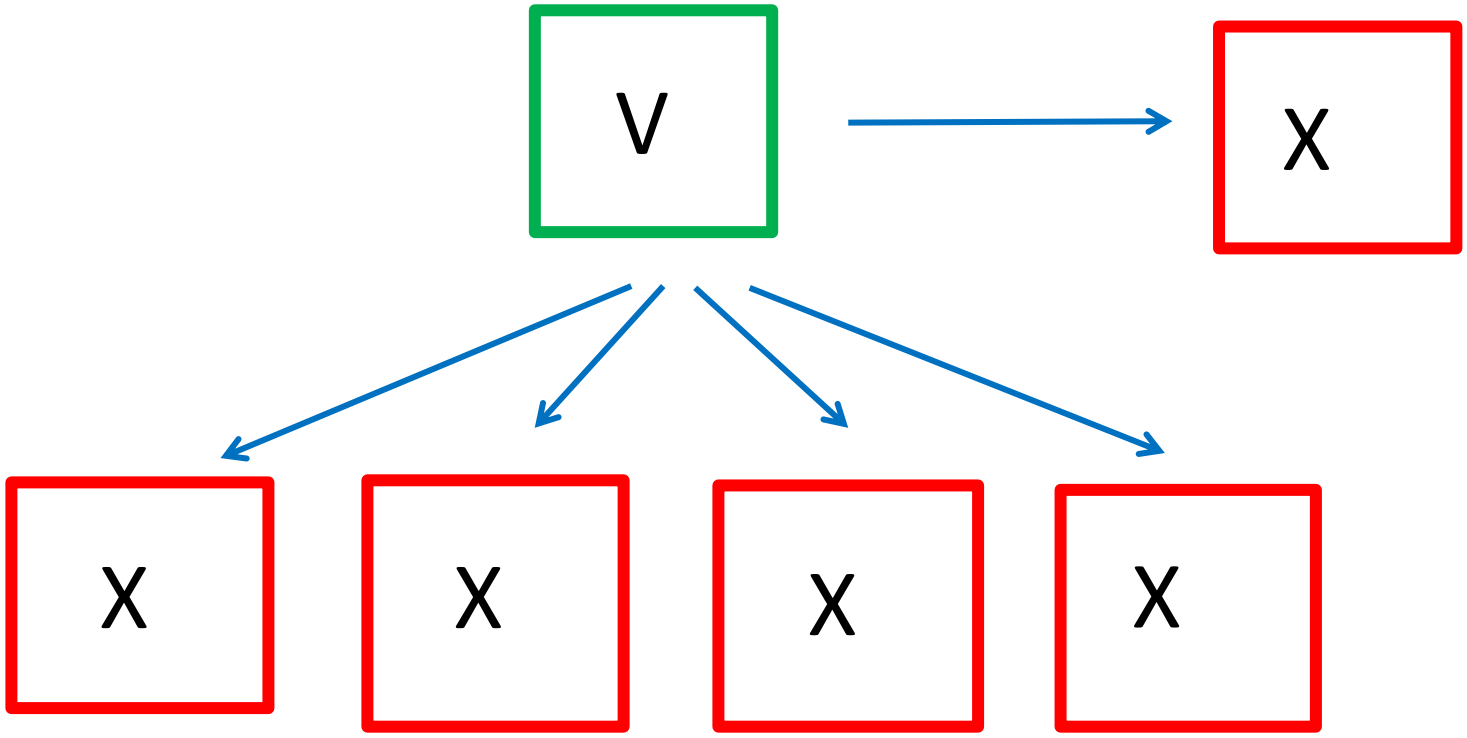
- Human studies
- Computational models
- Implications: representation for recognition, brain processing

Searching for Minimal Images



Over 15,000 subjects, laboratory controls

‘MIRC’ (Minimal Recognizable Configuration): all 5 descendants are unrecognizable



Sharp transition

Examples

Pairs

Parent – MIRC,
Child – ‘sub-MIRC’

0.93

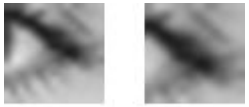
0.03

0.79

0.0

0.88

0.14



0.88 0.16

Plane

Ship

House fly

Bald eagle

Horse

Bike

Car door

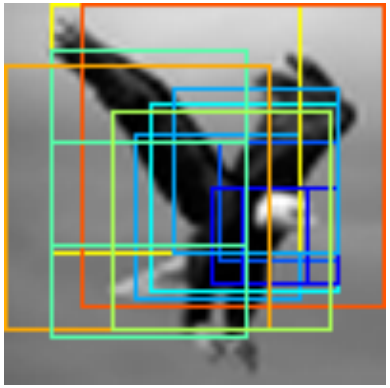
Human eye

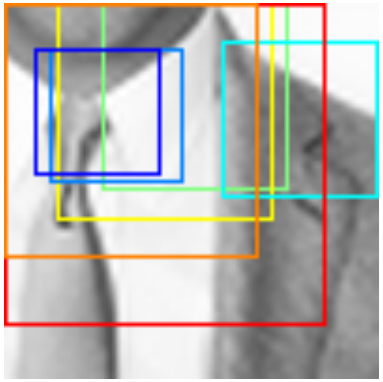
Eyeglasses

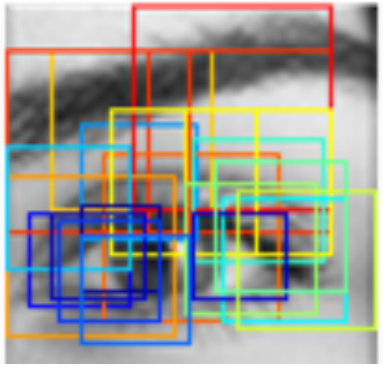
Suit&tie


0.60 14.63 px 46%
0.56 21.00 px 42%
0.52 16.00 px 32%
0.69 21.00 px 42%
0.58 20.00 px 44%
0.56 18.90 px 42%
0.68 14.80 px 40%
0.55 16.00 px 46%
0.56 18.29 px 57%
0.56 24.00 px 53%
0.74 16.20 px 60%
0.63 16.00 px 46%
0.72 16.00 px 46%
0.56 13.00 px 29%
0.54 25.00 px 56%
0.57 13.12 px 38%
0.66 18.00 px 60%
0.79 17.00 px 68%
0.67 26.00 px 52%
0.84 18.00 px 60%
0.57 13.00 px 26%
0.56 21.00 px 42%
0.63 21.00 px 42%
0.56 24.00 px 69%
0.53 20.00 px 50%
0.58 11.76 px 56%
0.56 16.20 px 60%
0.69 16.00 px 40%
0.65 17.14 px 57%
0.79 16.00 px 46%
0.87 15.43 px 57%
0.56 13.71 px 57%
0.54 11.70 px 26%
0.55 18.00 px 50%
0.60 17.31 px 42%
0.64 18.22 px 44%
0.68 18.90 px 42%
0.65 16.00 px 32%
0.52 13.00 px 26%
0.52 14.58 px 36%
0.73 14.00 px 47%
0.75 11.00 px 55%
0.59 16.57 px 57%
0.54 16.00 px 32%
0.72 20.00 px 44%
0.70 20.00 px 50%
0.79 30.00 px 67%
0.52 13.00 px 26%
0.61 13.00 px 26%
0.60 20.00 px 80%
0.79 20.00 px 44%
0.63 26.00 px 52%
0.74 18.29 px 57%
0.58 26.00 px 52%
0.56 15.30 px 34%
0.55 8.14 px 22%
0.93 20.00 px 40%
0.74 11.67 px 47%
0.52 28.80 px 64%
0.52 18.00 px 40%
0.74 8.00 px 27%
0.55 13.00 px 33%
0.61 16.00 px 40%
0.57 16.00 px 36%
0.55 11.00 px 55%
0.69 9.00 px 60%
0.89 8.00 px 32%
0.53 12.60 px 47%
0.65 14.86 px 57%
0.53 10.97 px 34%
0.52 14.50 px 50%
0.81 16.44 px 44%
0.75 16.44 px 44%
0.54 18.90 px 42%

Cover







Average 16.9 / class
Highly redundant

Each MIRC is non-redundant: each feature is important

- Sensitive tool to compare representations

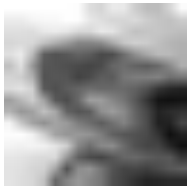
- Differences between MIRCs and sub-MIRCs to infer visual features

- Recognition features not captured by human feed-forward models and computer vision representations

Testing computational models

- Training of object images,
- Testing on minimal images
- MIRCs and sub-MIRCs

0.71



0.03

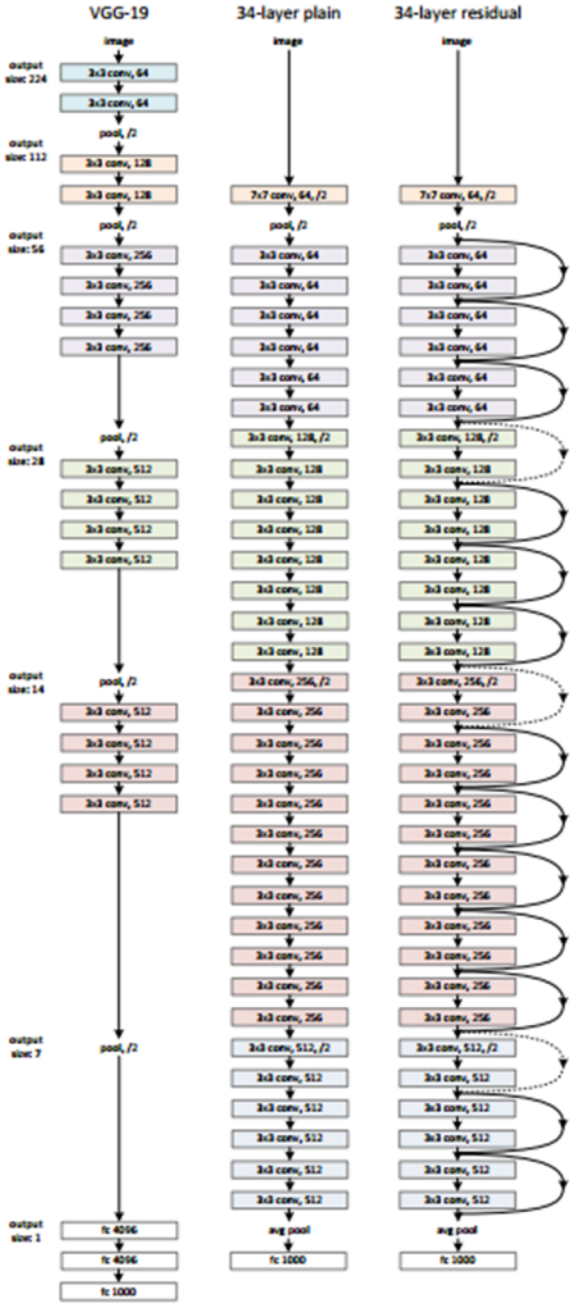
Recognition gap
Accuracy


CATHAY PACIFIC CARGO


AIR NEW ZEALAND

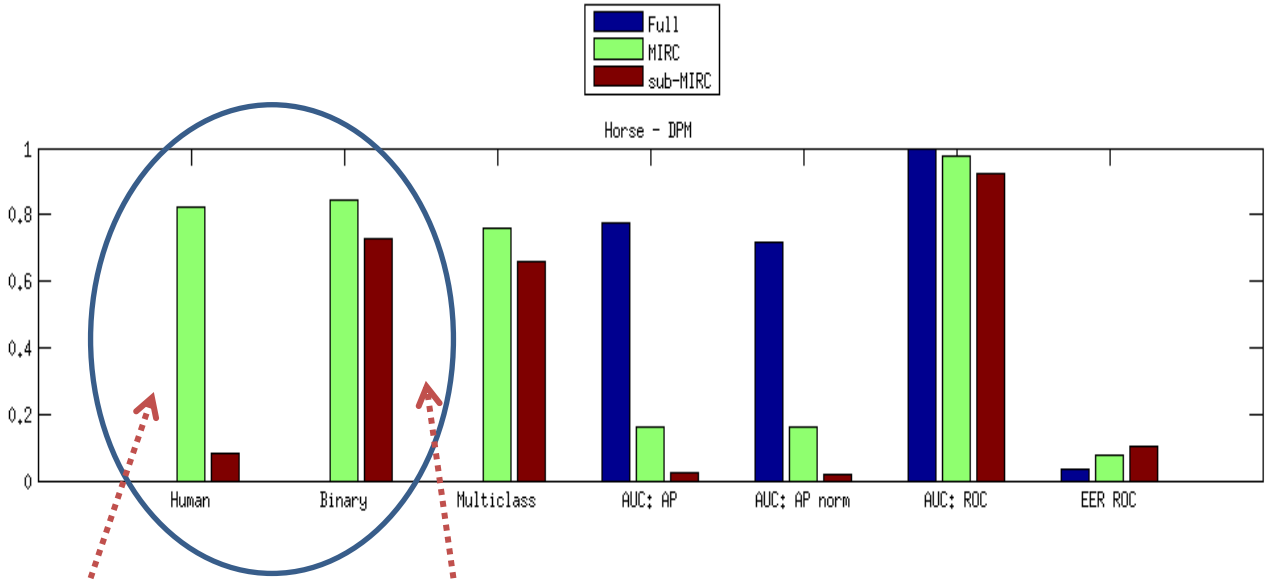

armavia


BOEING

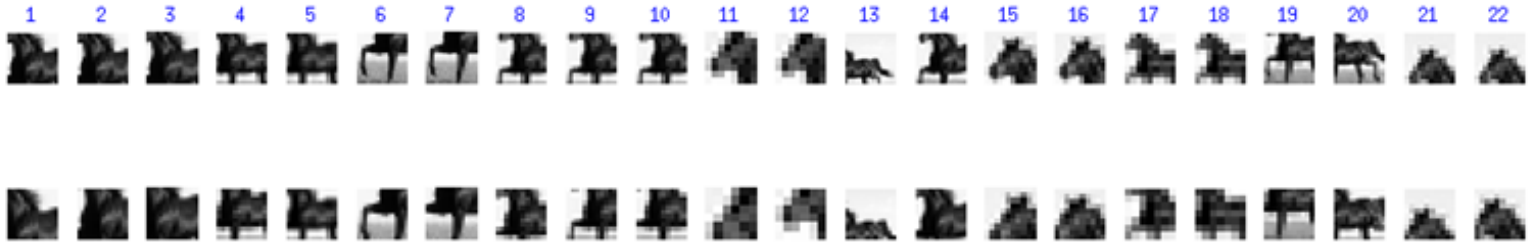


He et al Dec 2015

The recognition gap is not reproduced



Human Model



- Recognition of minimal images does not emerge by training any of the models tested.

- The large gap at the minimal level is not reproduced

- The accuracy of recognition is lower than human's

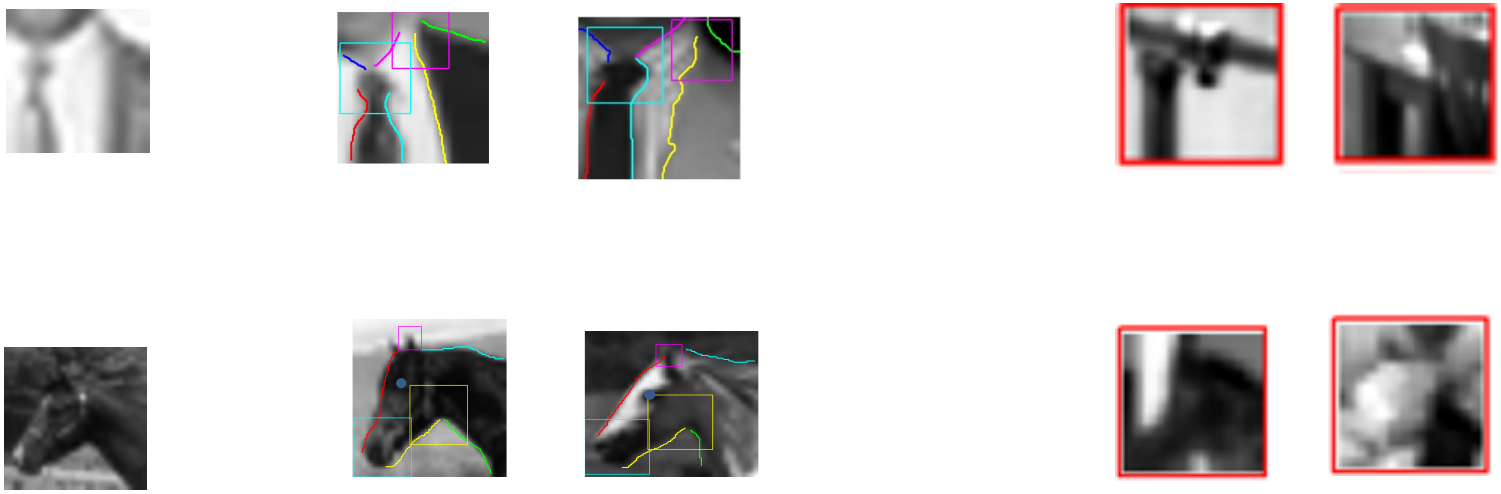
- Representations used by existing models do not capture differences that human recognition is sensitive to

MIRCs: Internal Interpretation

- Humans can interpret detailed sub-structures within the ‘atomic’ MIRC.
- Cannot be done by current feed-forward models

Internal Interpretation

Also for Validation

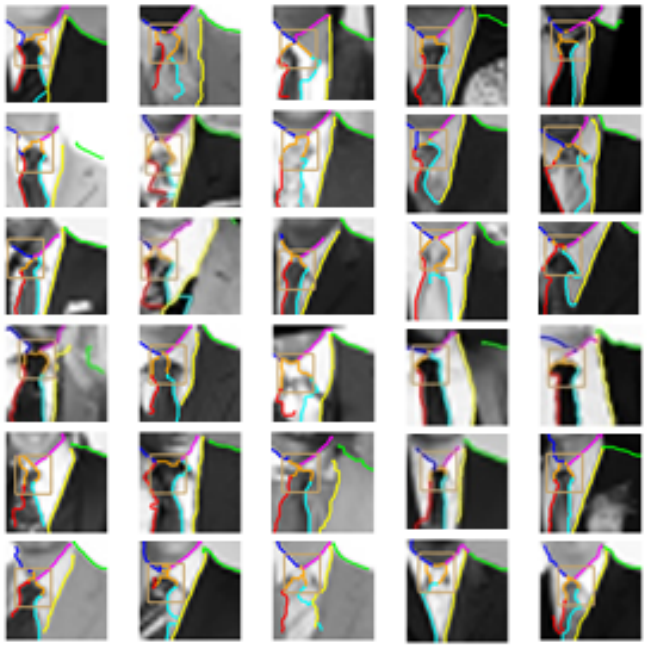


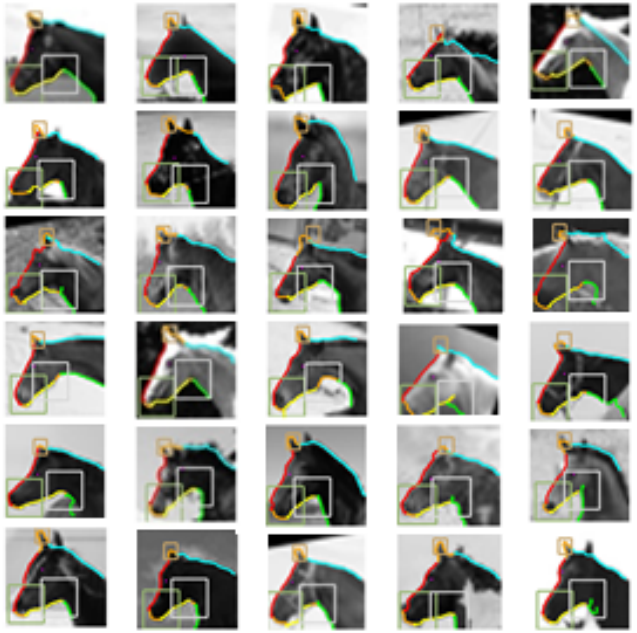
Internal interpretations, perceived by humans
Cannot be produced by existing feed-forward models

- Current recognition models are feed-forward only
- This is likely to limit their ability to providing interpretation of find details
- A model that can produce interpretation of MIRCs
- The model uses a top-down stage (back to V1)

Ben-Yossef et al, Cognition 2018
Full interpretation of minimal images

Results: Interpretation Model



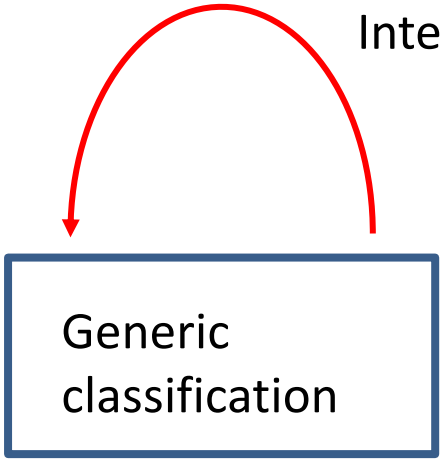


Two stages in recognition

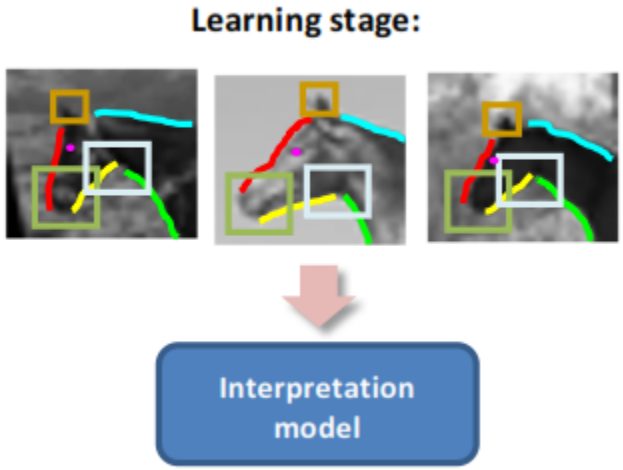
Class-specific validation
and interpretation
Internal and external

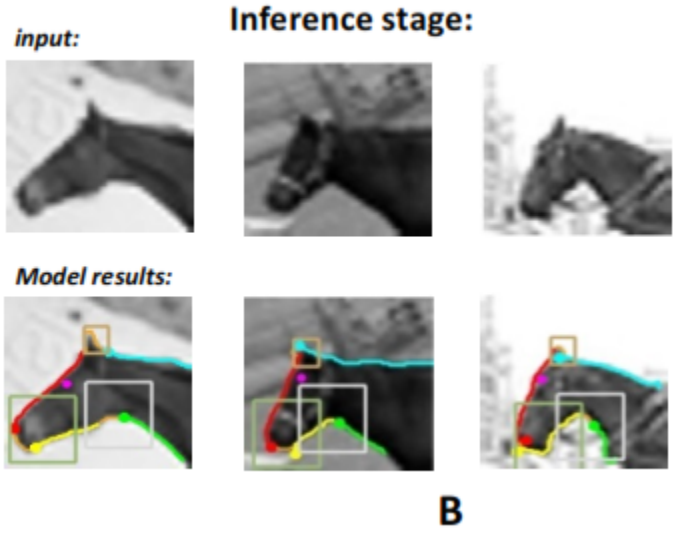
Entire image

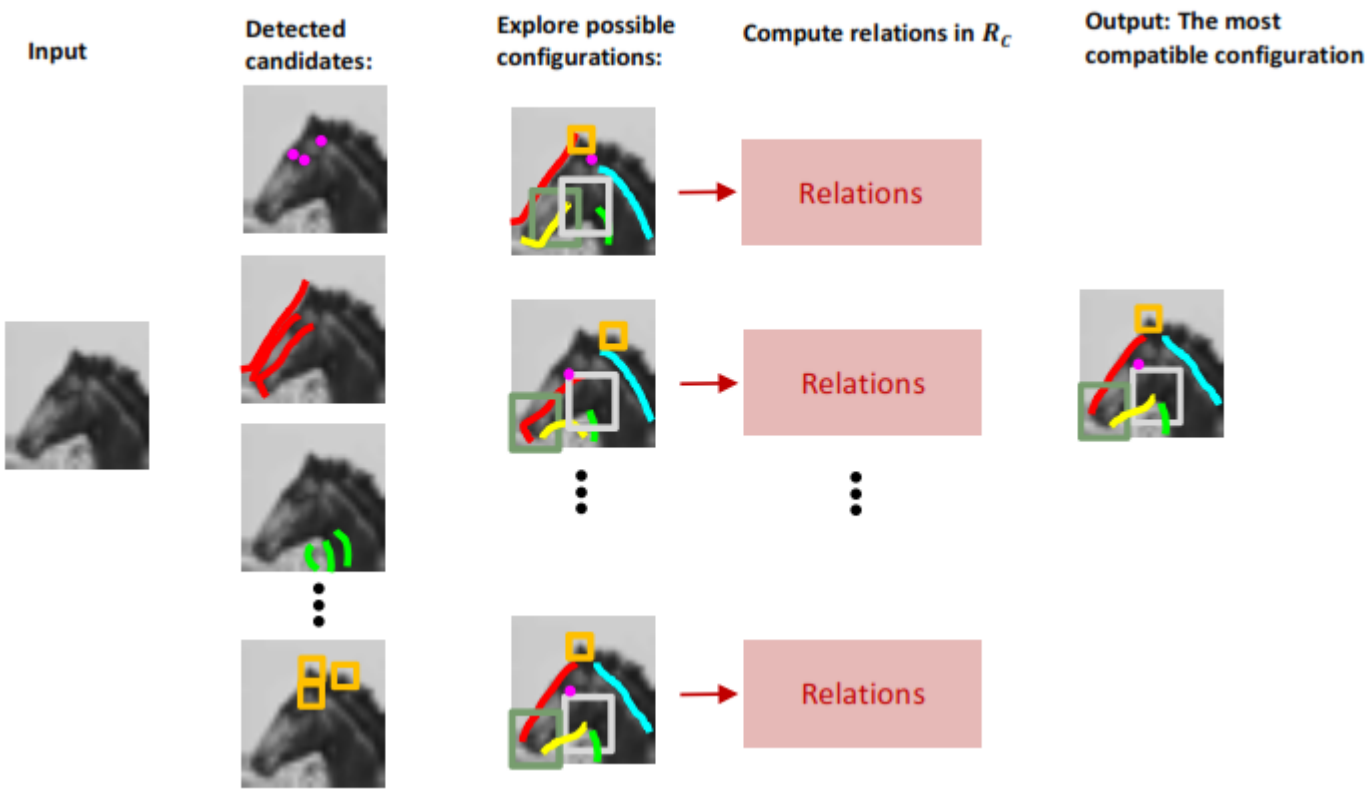
MIRCs level



Image


Learning stage:
Interpretation model
A


input:
Inference stage:
Model results:
B


Input
Detected candidates:
Explore possible configurations:
Compute relations in R_C
Output: The most compatible configuration
Relations
Relations
Relations

Psychophysics:
MIRC recognition takes longer than full images

- MIRC recognition as a function of exposure time
- Recognition of full images is often obtained with brief exposure
- MIRC recognition improves over long exposure time

Overall 277 Ss

Exposure time	Recognition rate
200 msec + mask (21 Ss)	**20%**
200 msec (32 Ss)	**44%**
500 msec + mask (16 Ss)	**42%**
500 msec (36 Ss)	**48%**
1000 msec + mask (21 Ss)	**51%**
1000 msec (49 Ss)	**58%**
2000msec +mask (19 Ss)	**58%**
2000 msec (47 Ss)	**59%**
Until response (36 Ss)	**68%**

Minimal images in human and computer vision

- Current models do not recognize well minimal images
- Current models do not produce full internal interpretation
- For humans MIRC recognition and detailed interpretation take longer than bottom-up recognition
- In human vision the size of the channels is about the same as a minimal image
- Future directions for study: MIRC recognition in the human bain

Possible project

- Applying available classification models to small images.
- Selecting small recognizable parts (small informal testing)
- Testing network recognition, gradually increasing patch size.