Emergence of transformation tolerant representations of visual objects in rat lateral extrastriate cortex

Davide Zoccolan
1. Background
   - Why vision?
   - Why rats (rodents)?
Why to study vision?

1. Because it’s an extraordinary complex cognitive function

   Sensation $\rightarrow$ Perception
Il riconoscimento visivo
Identificare/classificare gli oggetti in una scena
Il riconoscimento visivo
Identificare/classificare gli oggetti in una scena
Il riconoscimento visivo
Identificare/classificare gli oggetti in una scena
Il riconoscimento visivo: identificare/classificare gli oggetti in una scena.
Il riconoscimento visivo

Identificare/classificare gli oggetti in una scena

pixels = photoreceptors
Il riconoscimento visivo

Identificare/classificare gli oggetti in una scena

bananas
Il riconoscimento visivo
Identificare/classificare gli oggetti in una scena
clothes
Il riconoscimento visivo

Identificare/classificare gli oggetti in una scena

face
Why to study vision?

1. Because it’s an **extraordinary complex**
cognitive function

   Sensation $\rightarrow$ Perception
Why to study vision?

1. Because it’s an extraordinary complex cognitive function
   
   Sensation → Perception

2. Because the underlying cortical machinery is extraordinary complex
Why to study vision?
Why to study vision?

DiCarlo, Zoccolan and Rust (2012)
Why to study vision?

1. Because it’s an **extraordinary complex**
cognitive function
   
   Sensation → Perception

2. Because the underlying cortical machinery is **extraordinary complex**

Understanding vision:

→ understanding the brain
→ achieving artificial intelligence
1. Background
   - Why vision?
   - Why rats (rodents)?
Issue with monkey studies
Issue with monkey studies

the complexity of the primate visual system

... compared to ...

the limited experimental approaches
Is there any more manageable model system of visual functions?

1. Simpler visual system
   - lower number of visual areas
   - Less intricate feedforward and feedback circuitry

1. More experimentally accessible
   - high-throughput behavioral testing and large-scale recordings
   - molecular and genetic manipulation
   - In-vivo 2-photon imaging, optogenetics, etc.

2. Still embodying the core computations underlying visual object recognition and shape processing
Are rodents (mice and rats) suitable models to study higher-level vision? … largely neglected by vision scientists (until recently)
Can rats/mice see objects?

E.g., rats have a much poorer spatial resolution (~1 cycle/deg) compared to human fovea (~60 cycles/deg)

Since 2007 ... rodents are emerging as novel models of visual functions

**Mouse**
- Physiology/imaging of lower-level visual areas
  - 2007: Wang and Burkhalter (J Comp Neurol)
  - 2008: Niell and Stryker (J Neurosci)
  - 2009: Niell and Stryker (Neuron)
  - 2010: Marshel et al (Neuron), Anderman et al (Neuron)

**Rat**
- Psychophysics of visual perception/recognition
  - 2007: Greenberg et al (Nat Neurosci)
  - 2008: Zoccolan et al (PNAS)
  - 2009: Meyer et al (J Vision)
1. Background
   - Why vision?
   - Why rats (rodents)?

1. What is rat vision capable of?
   - Behavior
   - Neurophysiology
1. Background
   - Why vision?
   - Why rats (rodents)?

1. What is rat vision capable of?
   - Behavior
     • Zoccolan, Oertelt, DiCarlo, Cox (2009)
     • Tafazoli, Di Filippo, Zoccolan (2012)
     • Alemi-Neissi, Rosselli, Zoccolan (2013)
     • Rosselli, Alemi-Neissi, Ansuini, Zoccolan (2015)
     • Djurdjevic, Bertolini, Ansuini, Zoccolan (in preparation)
   - Neurophysiology
1. Background
   - Why vision?
   - Why rats (rodents)?

1. What is rat vision capable of?
   - Behavior
     - Zoccolan, Oertelt, DiCarlo, Cox (2009)
     - Tafazoli, Di Filippo, Zoccolan (2012)
     - Alemi-Neissi, Rosselli, Zoccolan (2013)
     - Rosselli, Alemi-Neissi, Ansuini, Zoccolan (2015)
     - Djurdjevic, Bertolini, Ansuini, Zoccolan (in preparation)
   - Neurophysiology
     - Tafazoli, Safaai, De Franceschi, Rosselli, Riggi, Buffolo, Panzeri, Zoccolan (2017)
Summary of the behavioral studies

1. Rats **are capable** of invariant visual object recognition
   (Zoccolan, Oertelt, DiCarlo, Cox, 2009)

   - Frontal views (0° in-depth rotation)
   - Default size (40° of visual angle)
   - Default position (center of the screen)
Summary of the behavioral studies

1. Rats are capable of invariant visual object recognition (Zoccolan, Oertelt, DiCarlo, Cox, 2009)
Summary of the behavioral studies

1. Rats are capable of invariant visual object recognition
   (Zoccolan, Oertelt, DiCarlo, Cox, 2009)
Summary of the behavioral studies

1. Rats **are capable** of invariant visual object recognition
   
   *(Zoccolan, Oertelt, DiCarlo, Cox, 2009)*

---

**Diagram:**

- **Size (°):** 15, 20, 25, 30, 35, 40
- **Rotation in depth (°):** -60, -45, -30, -15, 0, 15, 30, 45, 60

- Two sets of objects are depicted, each showing different combinations of size and rotation in depth.
Summary of the behavioral studies

1. Rats are capable of invariant visual object recognition (Zoccolan, Oertelt, DiCarlo, Cox, 2009)

108 size-rotation combinations
The rat in action
Summary of the behavioral studies

1. Rats are capable of invariant visual object recognition *(Zoccolan, Oertelt, DiCarlo, Cox, 2009)*

<table>
<thead>
<tr>
<th>Size (°)</th>
<th>15</th>
<th>20</th>
<th>25</th>
<th>30</th>
<th>35</th>
<th>40</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rotation in depth (°)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-60</td>
<td>*</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
</tr>
<tr>
<td>-45</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
</tr>
<tr>
<td>-30</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
</tr>
<tr>
<td>-15</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
</tr>
<tr>
<td>0</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
</tr>
<tr>
<td>15</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
</tr>
<tr>
<td>30</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
</tr>
<tr>
<td>45</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
</tr>
<tr>
<td>60</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
</tr>
</tbody>
</table>

Significance level: (*p < 0.05; **p < 0.01; ***p < 0.001)
Summary of the behavioral studies

1. Rats are capable of invariant visual object recognition (Zoccolan, Oertelt, DiCarlo, Cox, 2009)
Rats are capable of invariant visual object recognition (Zoccolan, Oertelt, DiCarlo, Cox, 2009)

Rats spontaneously generalize their recognition to novel object views (Tafazoli, Di Filippo, Zoccolan, 2012)

Default view

perceptually similar? → Visual priming paradigm

16 novel views
Summary of the behavioral studies

1. Rats are capable of invariant visual object recognition (Zoccolan, Oertelt, DiCarlo, Cox, 2009)

1. Rats spontaneously generalize their recognition to novel object views (Tafazoli, Di Filippo, Zoccolan, 2012)

2. Rats rely on an invariant, shape-based, multifeatural processing strategy of visual objects (Alemi-Neissi, Rosselli, Zoccolan, 2013)
Summary of the behavioral studies

1. Rats are capable of invariant visual object recognition (Zoccolan, Oertelt, DiCarlo, Cox, 2009)

1. Rats spontaneously generalize their recognition to novel object views (Tafazoli, Di Filippo, Zoccolan, 2012)

2. Rats rely on an invariant, shape-based, multifeatural processing strategy of visual objects (Alemi-Neissi, Rosselli, Zoccolan, 2013)

Rats → good models to study the neuronal basis of visual object recognition and shape processing
Layout of the talk

1. Background
   - Why vision?
   - Why rats (rodents)?

1. What is rat vision capable of?
   - Behavior
   - Neurophysiology
Does an homologous of the primate **ventral stream** exist in the rat brain?
The primate ventral stream: an object processing pathway

Adapted from Rousselet et al. (2004)
The primate ventral stream: an object processing pathway

Adapted from Rousselet et al. (2004)

Adapted from Kandel, Schwartz and Jessell
The primate ventral stream: an object processing pathway

Adapted from Rousselet et al. (2004)

Kobatake et al. (1994)
The primate ventral stream: an object processing pathway

Adapted from Rousselet et al. (2004)
The primate ventral stream: an object processing pathway

Adapted from Rousselet et al. (2004)
The primate ventral stream: builds explicit representations of visual objects
Does an homologous of the primate ventral stream exist in the rat brain?
Hierarchical organization of rat visual areas

Adapted from Sereno & Allaman (1991)

Adapted from Coogan & Burkhalter (1993)
Hierarchical organization of rat visual areas → the **object processing** pathway

Adapted from Sereno & Allaman (1991)

Adapted from Coogan & Burkhalter (1993)
Experimental design
1. multi-electrode recordings in anesthetized rats
Experimental design

2. stimulus set and presentation

visual stimuli

object:  

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>#1</td>
<td>#2</td>
<td>#3</td>
<td>#4</td>
<td>#5</td>
</tr>
<tr>
<td>#6</td>
<td>#7</td>
<td>#8</td>
<td>#9</td>
<td>#10</td>
</tr>
</tbody>
</table>
Experimental design

2. stimulus set and presentation

**visual stimuli**

<table>
<thead>
<tr>
<th>object:</th>
<th>#6</th>
<th>#7</th>
<th>#8</th>
<th>#9</th>
<th>#10</th>
</tr>
</thead>
</table>

**identity-preserving transformations**

1. **position changes**
   (size = 35°; in-plane & in-depth rotations = 0°; luminance = 100%)

<table>
<thead>
<tr>
<th>-22.5°</th>
<th>-15°</th>
<th>-7.5°</th>
<th>0°</th>
</tr>
</thead>
<tbody>
<tr>
<td>![Image of object #6 at -22.5°]</td>
<td>![Image of object #7 at -15°]</td>
<td>![Image of object #6 at -7.5°]</td>
<td>![Image of object #6 at 0°]</td>
</tr>
<tr>
<td>+7.5°</td>
<td>+15°</td>
<td>+22.5°</td>
<td>+30°</td>
</tr>
<tr>
<td>![Image of object #6 at +7.5°]</td>
<td>![Image of object #7 at +15°]</td>
<td>![Image of object #6 at +22.5°]</td>
<td>![Image of object #6 at +30°]</td>
</tr>
</tbody>
</table>
Experimental design

2. **stimulus set and presentation**

RSVP: Rapid Sequence Visual Presentation

![Image showing a sequence of images with 250 ms intervals](image-url)
Experimental design

3. receptive fields mapping

10° long moving bars presented over a 6x11 grid
Results: basic properties

1. **receptive field size** increases from V1 to LL
Results: basic properties
1. **receptive field size** increases from V1 to LL
Results: basic properties

1. **receptive field size** increases from V1 to LL
Results: basic properties

1. **receptive field size** increases from V1 to LL
Results: basic properties

1. **receptive field size** increases from V1 to LL
Results: basic properties

1. **receptive field size** increases from V1 to LL

![Graph showing receptive field size increases from V1 to LL](chart.png)
Results: basic properties

1. **receptive field size** increases from V1 to LL

![Graph showing receptive field size across different visual cortical areas](image-url)
Results: basic properties

1. **receptive field size** increases from V1 to LL

3. **latency** increases from V1 to LL
Results: basic properties
1. **receptive field size** increases from V1 to LL
2. **latency** increases from V1 to LL
Results: basic properties

1. **receptive field size** increases from V1 to LL
2. **latency** increases from V1 to LL

The monkey ventral stream

- **Latency**
  - AIT: 80-100 ms
  - PIT: 70-90 ms
  - V4: 60-80 ms
  - V2: 50-70 ms
  - V1: 40-60 ms

- **RF size**
  - AIT: 2.5°-70°
  - PIT: 2°-25°
  - V4: 1°-20°
  - V2: 0.5°-4°
  - V1: 0.5°-1.5°

Adapted from Rousselet et al. (2004)
Results: basic properties

1. **receptive field size** increases from V1 to LL
2. **latency** increases from V1 to LL

The monkey ventral stream

- **Tuning complexity**
- **Latency**
- **RF size**

<table>
<thead>
<tr>
<th></th>
<th>AIT</th>
<th>80-100 ms</th>
<th>2.5°-70°</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PIT</td>
<td>70–90 ms</td>
<td>2°–25°</td>
</tr>
<tr>
<td></td>
<td>V4</td>
<td>60–80 ms</td>
<td>1°–20°</td>
</tr>
<tr>
<td></td>
<td>V2</td>
<td>50–70 ms</td>
<td>0.5°–4°</td>
</tr>
<tr>
<td></td>
<td>V1</td>
<td>40-60 ms</td>
<td>0.5°–1.5°</td>
</tr>
</tbody>
</table>

Adapted from Rousselet et al. (2004)
Shape processing & invariant recognition

Low-level representation

V1 representation?

High-level representation

LL representation?

LL representation?

“car”
The approach:
1. **mutual information** between stimulus and response
2. **decoding** stimulus identity using machine learning

*Quiroga & Panzeri (2009)*
Examples of neuronal tuning

V1 neuron

very effective object

poorly effective object
Examples of neuronal tuning

V1 neuron

LL neuron
Examples of neuronal tuning

Two sources of modulation of neuronal firing rate:

1. variation in **object transformation** $T$
2. variation in **object identity** $O$
First question: **total information per neuron** across visual areas

- How much **information** do neurons carry about the **identity** of all our **object views**?

→ all object views treated as independent stimuli

<table>
<thead>
<tr>
<th>stimulus:</th>
<th>vs</th>
<th>vs</th>
<th>vs</th>
<th>vs</th>
</tr>
</thead>
<tbody>
<tr>
<td>[Image]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[Image]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[Image]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[Image]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Probability

Response (firing rate)
First question: total information per neuron across visual areas

- How much information do neurons carry about the identity of all our object views?

→ all object views treated as independent stimuli

stimulus: vs vs vs vs

response:
First question: \textbf{total information per neuron} across visual areas

- How much \textbf{information} do neurons carry about the identity of all our object views?

\textbf{Approach:}

\rightarrow \text{Compute} \textbf{mutual information} between stimulus and response:

\[ I(R; S) = \sum_s P(s) \sum_r P(r/s) \log_2 \frac{P(r/s)}{P(r)} \]

\( S = \text{stimulus} = O \& T \)
Mutual information: **results**

1. **stimulus information** across visual areas
Mutual information: results
1. stimulus information across visual areas

[Bar chart showing information (bits) for different visual cortical areas: V1, LM, LI, LL. The bars are labeled with numbers (227, 130, 260, 152) and error bars are present. Asterisks indicate significance levels: *** for V1, ** for LM, *** for LI, and no asterisk for LL.]

[Brain diagram with various visual areas labeled: V1, LM, LI, LL, AL, RL, AM, M, P1, P2, L1, L1a, PL, L1, L1a.]
Mutual information: **conclusions**

1. Mutual information between neuronal response and individual stimulus conditions decreases from V1 to LL
Mutual information: conclusions

1. Mutual information between neuronal response and individual stimulus conditions decreases from V1 to LL

   ➢ consistent with the existence of a processing hierarchy
Second question: what kind of information do neurons code?

“lowest-level” visual attributes

- overall luminance/contrast
  - low luminance
  - high luminance

“higher-order” features

- oriented edges & corners
- luminance patches
- curved boundaries
Variation of luminous energy could explain:

1. neuronal sensitivity to **object transformations**
2. neuronal selectivity for **object identity**

sensitivity to **transformation**
Effective RF luminance:
the luminous energy impinged by a stimulus on a neuronal RF
Dependence of neuronal firing on RF luminance
Dependence of neuronal firing on RF luminance

Take every stimulus $S_i$ (10 objects x 23 transformations) and:

1. Compute the corresponding RF luminance value

2. Bin the resulting luminance axis in $N$ equipopulated bins
Dependence of neuronal firing on RF luminance

V1 neuron
Dependence of neuronal firing on RF luminance

**LL neuron**

**V1 neuron**
Dependence of neuronal firing on RF luminance

**LL neuron**

- **Firing Rate (spikes/s)**
- **Position (°)**

**V1 neuron**

- **Firing Rate (spikes/s)**
- **Position (°)**

**Ranked objects**

- **RF luminance (bin #)**
- **Firing rate (spikes/s)**

*Low luminance*

*High luminance*
Second question:
**what kind of information** do neurons code?

- What fraction of information do neurons carry about the **overall luminance/contrast** of visual objects?
- What fraction of information do neurons carry about **higher-order features**?
  - higher-order = everything not accounted by luminance alone
Second question: what **kind of information** do neurons code?

- What fraction of information do neurons carry about the **overall luminance/contrast** of visual objects?
- What fraction of information do neurons carry about **higher-order features**?

\[
I(R; S) = I(R; S' \& L)
\]

\[
L = \text{RF luminance}
\]

\[
S' = \text{object identity defined by higher-order features}
\]

\[
S = \text{stimulus} = S' \& L
\]
Second question: what kind of information do neurons code?

- What fraction of information do neurons carry about the overall luminance/contrast of visual objects?
- What fraction of information do neurons carry about higher-order features?

\[
I(R; S) = I(R; S'&L) = I(R; S'|L) + I(R; L)
\]

\(L = \text{RF luminance}\)

\(S' = \text{object identity defined by higher-order features}\)

\(S = \text{stimulus} = S' \& L\)

\(\text{higher-order information}\)

\(\text{luminance information}\)
Mutual information: results

2. fraction of higher-order information

\[ I(R; S' & L) = I(R; S'|L) + I(R; L) \]

**Information (bits)**

- V1: 227
- LM: 130
- LI: 260
- LL: 152

Visual cortical area

**Total information**
Mutual information: results
2. fraction of higher-order information

\[ I(R; S'\&L) = I(R; S'|L) + I(R; L) \]

- **I(R; S'|L)**: higher-order information
- **I(R; L)**: luminance information

<table>
<thead>
<tr>
<th>Visual cortical area</th>
<th>Information (bits)</th>
</tr>
</thead>
<tbody>
<tr>
<td>V1</td>
<td>227</td>
</tr>
<tr>
<td>LM</td>
<td>130</td>
</tr>
<tr>
<td>LI</td>
<td>260</td>
</tr>
<tr>
<td>LL</td>
<td>152</td>
</tr>
</tbody>
</table>
Mutual information: **results**

2. fraction of **higher-order** information

Fraction of information about object **higher-order features**: 

\[ f_{\text{high}} = \frac{I(R; S' | L)}{I(R; S' \& L)} \]
Mutual information: **conclusions**

1. Mutual information between neuronal response and individual stimulus conditions decreases from V1 to LL
   ➢ consistent with the existence of a processing hierarchy

2. The fraction of **information** carried by neurons about higher-order features **increases** from V1 to LL
   ➢ consistent with formation of higher-order representations
Third question: how **view-invariant** is object information?

- Take a pair of objects: \( O = \{o_i, o_j\} \)
- Take multiple transformations per object: \( T = \{t_1, \ldots, t_{23}\} \)
- Define each stimulus (object view) as: \( S = O \& T \)
- Decompose the total stimulus information as:

\[
I(R; S) = I(R; O&T)
\]

**total information**
Third question: how **view-invariant** is object information?

- Take a pair of objects: \( O = \{ o_i , o_j \} \)
- Take multiple transformations per object: \( T = \{ t_1 , \ldots , t_{23} \} \)
- Define each stimulus (object view) as: \( S = O \& T \)
- Decompose the total stimulus information as:

\[
I(R; S) = I(R; O&T) = I(R; T|O) + I(R; O)
\]

- **Total information**
- **View-specific information**
- **View-invariant information**
Third question: how **view-invariant** is object information?

\[ I(R; S) = I(R; O&T) = I(R; T|O) + I(R; O) \]

(total information)

(view-invariant information)
Mutual information: **results**

3. **view-invariant** object information

**view-invariant information**

```
<table>
<thead>
<tr>
<th>Visual cortical area</th>
<th>Information per neuron (bits)</th>
</tr>
</thead>
<tbody>
<tr>
<td>V1</td>
<td>0.001</td>
</tr>
<tr>
<td>LM</td>
<td>0.001</td>
</tr>
<tr>
<td>LI</td>
<td>0.002</td>
</tr>
<tr>
<td>LL</td>
<td>0.003</td>
</tr>
</tbody>
</table>
```

V1: 228, LM: 131, LI: 260, LL: 152
Mutual information: results
3. view-invariant object information

View-invariant object information per neuron (bits)

<table>
<thead>
<tr>
<th>Area</th>
<th>V1</th>
<th>LM</th>
<th>LI</th>
<th>LL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Value</td>
<td>228</td>
<td>131</td>
<td>260</td>
<td>152</td>
</tr>
</tbody>
</table>

Fraction of view-invariant object information per neuron (%)

<table>
<thead>
<tr>
<th>Area</th>
<th>V1</th>
<th>LM</th>
<th>LI</th>
<th>LL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Value</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Mutual information: **conclusions**

1. Mutual information between neuronal response and individual stimulus conditions decreases from V1 to LL.

2. The fraction of information carried by neurons about higher-order features increases from V1 to LL.

3. The **view-invariant object information per neuron** increases from V1 to LL.

   ➢ consistent with **formation of invariant representations**
The approach:

1. **mutual information** between stimulus and response
2. **decoding** stimulus identity using machine learning
The approach:

1. **mutual information** between stimulus and response
2. **decoding** stimulus identity using machine learning

*Quiroga & Panzeri* (2009)

**mutual information** → **upper bound** to how well we can reconstruct a stimulus property (e.g., object luminance or identity) from observing a neuronal response.
The approach:
1. **mutual information** between stimulus and response
2. **decoding** stimulus identity using machine learning

**decoding** → how easily a stimulus property (e.g., object identity) can be read out from the neuronal response?

**mutual information** → **upper bound** to how well we can reconstruct a stimulus property (e.g., object luminance or identity) from observing a neuronal response

Quiroga & Panzeri (2009)
Forth question: how **easily readable** is object identity?

large view-invariant information & *low* linear separability
Forth question: how easily readable is object identity?

large view-invariant information & \textit{low} linear separability

large view-invariant information & \textit{large} linear separability
Forth question: how **easily readable** is object identity?

- Take a pair of objects
- Take multiple views of each object
- Key question:
Forth question: how **easily readable** is object identity?

- Take a pair of objects
- Take multiple views of each object
- **Key question:**
  - **If:** a decoder learns to discriminate two views of these objects
  - **then:** does it generalize to any other pair of views?
  
  → **Goodness of a representation to support invariant recognition**
Good vs. bad representations: how **easily readable** is object identity?
Good vs. bad representations: how **easily readable** is object identity?
Good vs. bad representations: how **easily readable** is object identity?
Good vs. bad representations: how *easily readable* is object identity?
Good vs. bad representations: how \textbf{easily readable} is object identity?
Good vs. bad representations: how **easily readable** is object identity?
Good vs. bad representations: how **easily readable** is object identity?
Method: measure **generalization** of recognition performance
Single-cell decoding: results
generalization to novel object views
Single-cell decoding: **results**
generalization to novel object views

![Graph showing single-neuron performance on a generalization task (bits)]

- V1: 228
- LM: 131

Visual cortical area

Brain diagram with labels: V1, LM, LI, LL, etc.
Single-cell decoding: results
generalization to novel object views
Single-cell decoding: results
generalization to novel object views
Single-cell decoding: results
generalization to novel object views

---

**Graph:**

- **x-axis:** Visual cortical area
- **y-axis:** Single-neuron performance on a generalization task (bits)

- **V1:** 228 bits
- **LM:** 131 bits
- **LI:** 260 bits
- **LL:** 152 bits

Significant differences indicated by ***p < 0.001***
Decoding: conclusions

1. Coding of object identity becomes more invariant to transformation from V1 to LL

➢ consistent with formation of invariant representations
Can we observe **macroscopic differences** at the neuronal population level?
Last question: *invariant recognition* at the neuronal population level
Last question:
**invariant recognition** at the neuronal population level
Last question: invariant recognition at the neuronal population level
Last question: invariant recognition at the neuronal population level
Last question: invariant recognition at the neuronal population level
Last question: **invariant recognition at the neuronal population level**
Last question:
invariant recognition at the neuronal population level
Last question: invariant recognition at the neuronal population level

Performance on a generalization task (bits)

Number of neurons

$n = 3$ (object pairs)
Last question: invariant recognition at the neuronal population level
Decoding: conclusions

1. Coding of object identity becomes more invariant to transformation from V1 to LL
   - At both the single-cell and neuronal population level
   - Consistent with formation of invariant representations

![Graph showing performance on a generalization task](image)
Rats are interesting models of object vision.

**Behavior**

Perceived similarity (priming magnitude) across different conditions.

**Neurophysiology**

Information (bits) and performance on a generalization task (bits) across visual cortical areas (V1, LM, LI, LL).
Acknowledgements

Sina Tafazoli  
Gioa De Franceschi

Alessandro Di Filippo  
Margherita Riggi

Alireza Alemi-Neissi  
Federica Buffolo

David Cox  
Nadja Oertelt

James DiCarlo  
Houman Safaai

Bold font = first authors or co-authors
Acknowledgements
Experimental design

4. identification of visual cortical areas

Base: ch. 32
Tip: ch. 1

recording area (32 sites)
Experimental design

4. identification of visual cortical areas

RF maps

Nasal
Temporal

Base: ch. 32
Tip: ch. 1

recording area (32 sites)
Experimental design
4. identification of **visual cortical areas**

![Diagram](image)

- **Base:** ch. 32
- **Tip:** ch. 1

**RF maps**
- Nasal: RF centers
- Temporal: RF centers

**Recording area (32 sites)**
Experimental design
4. identification of visual cortical areas

Base: ch. 32
Tip: ch. 1

RF maps
Nasal
Temporal
30 29 27 23 21
9 13 15 18 20
8 6 5 3 1

recording area (32 sites)
Experimental design

4. identification of visual cortical areas

RF maps

RF centers

recording area (32 sites)

electrode array

Base: ch. 32

Tip: ch. 1

Nasal

Temporal

10°
Experimental design

4. identification of visual cortical areas

Base: ch. 32
Tip: ch. 1

RF centers

RF maps

Nasal
Temporal

10°
The confound of luminance tuning
The confound of luminance tuning

V1 neuron

LL neuron
The confound of luminance tuning

V1 neuron

Position (°)

Firing Rate (spikes/s)

LL neuron

Position (°)

Firing Rate (spikes/s)
The confound of luminance tuning

V1 neuron

LL neuron
The confound of luminance tuning

V1 neuron

LL neuron
The confound of luminance tuning

“apparent” invariance

“real” invariance

V1 neuron

LL neuron
Decoding: **results**

(after matching RF luminance)
Decoding: results
(after matching RF luminance)
Decoding: results

(after matching RF luminance)
Decoding: results
(after matching RF luminance)
Decoding: results
(after matching RF luminance)

[Bar chart showing discrimination performance above chance level across different visual cortical areas (V1, LM, LI, LL) with corresponding data points and error bars.]
Decoding: results
(after matching RF luminance)
Decoding: results
(after matching RF luminance)
Decoding: results
(after matching RF luminance)
Decoding: conclusions

1. Coding of object identity becomes more invariant to transformation from V1 to LL
   - consistent with formation of invariant representations
Decoding: results

2. superficial vs. deep cortical layers
Decoding: results

2. superficial vs. deep cortical layers

recordings from deep layers
Decoding: results

2. superficial vs. deep cortical layers

recordings from deep layers

recordings from superficial layers
Decoding: results

2. superficial vs. deep cortical layers

![Graph showing discrimination performance above chance level for superficial layers (II - IV) and deep layers (V - VI).](image)
Decoding: results

3. individual variation axes
Decoding: **results**

3. **individual** variation axes

![Graph](Image)
Decoding: results
3. individual variation axes
Decoding: results

4. **parametric** position changes

![Graph showing the relationship between distance and discrimination performance. The x-axis represents the distance between training and testing positions in degrees, ranging from 7.5 to 30. The y-axis represents the percentage above chance level of discrimination performance. The graph includes lines for LL, LI, LM, and V1, with different colors and shaded areas indicating variability.]
Decoding: conclusions

1. Mutual information between neuronal response and individual stimulus conditions decreases from V1 to LL.

2. The fraction of information carried by neurons about "shape" attributes increases from V1 to LL.

3. Coding of object identity becomes more invariant to transformation from V1 to LL.

➢ consistent with formation of invariant representations.