

The human side of computer vision

Olga Russakovsky



My research: Computational building blocks of visual intelligence



My research: Computational building blocks of visual intelligence

Object detection



My research: Computational building blocks of visual intelligence

Object detection

Scene understanding



My research: Computational building blocks of visual intelligence

Object detection

Scene understanding

Action recognition

...



Standard computer vision paradigm

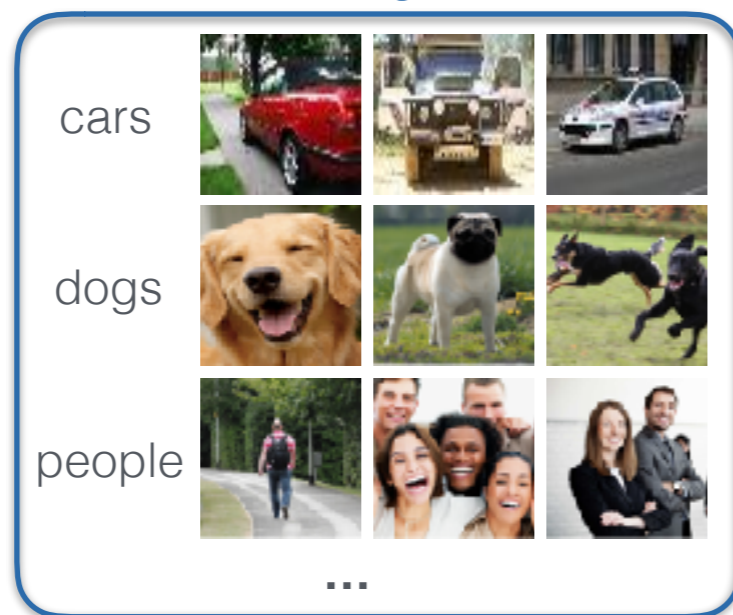
What objects are in the image?



Standard computer vision paradigm

What objects are in the image?

Training data

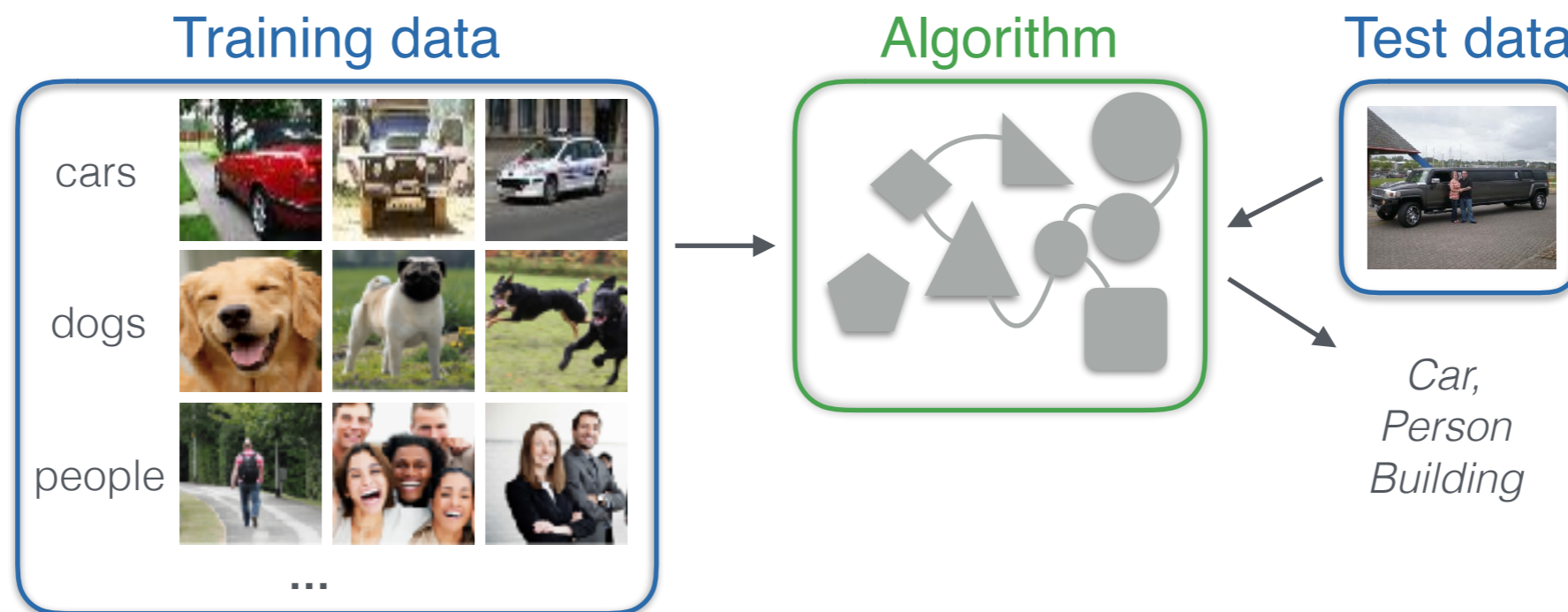


Test data



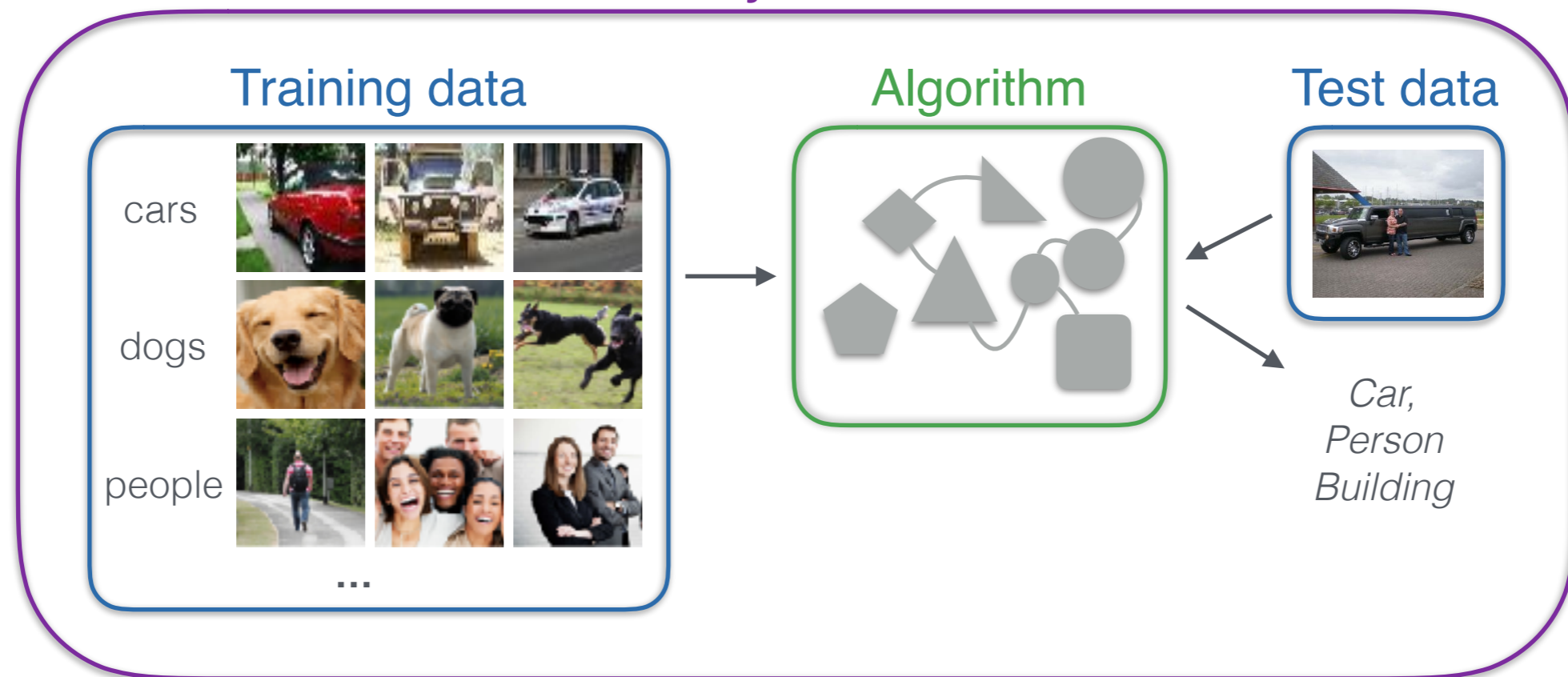
Standard computer vision paradigm

What objects are in the image?



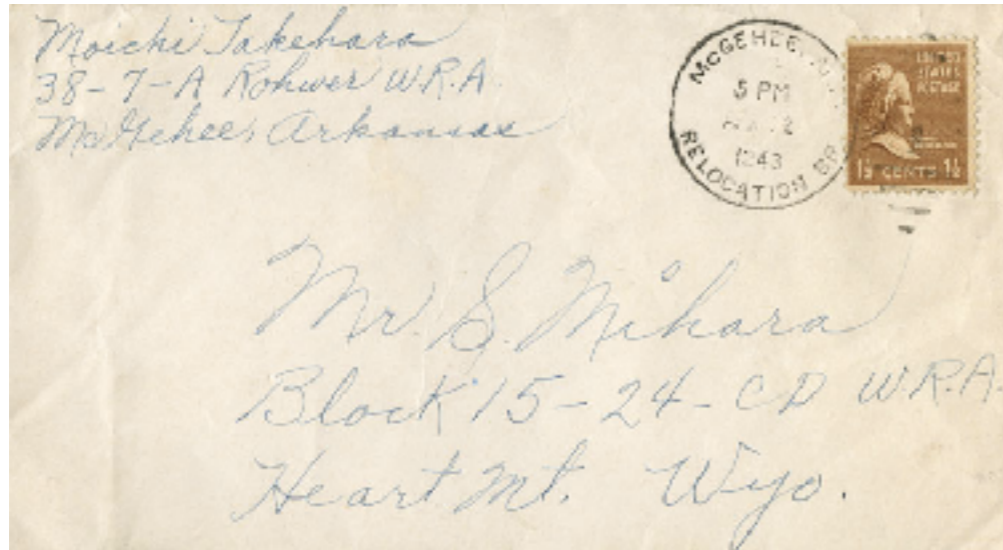
Standard computer vision paradigm

AI vision system



Success stories

Sorting mail



Detecting (frontal) faces



Depositing checks

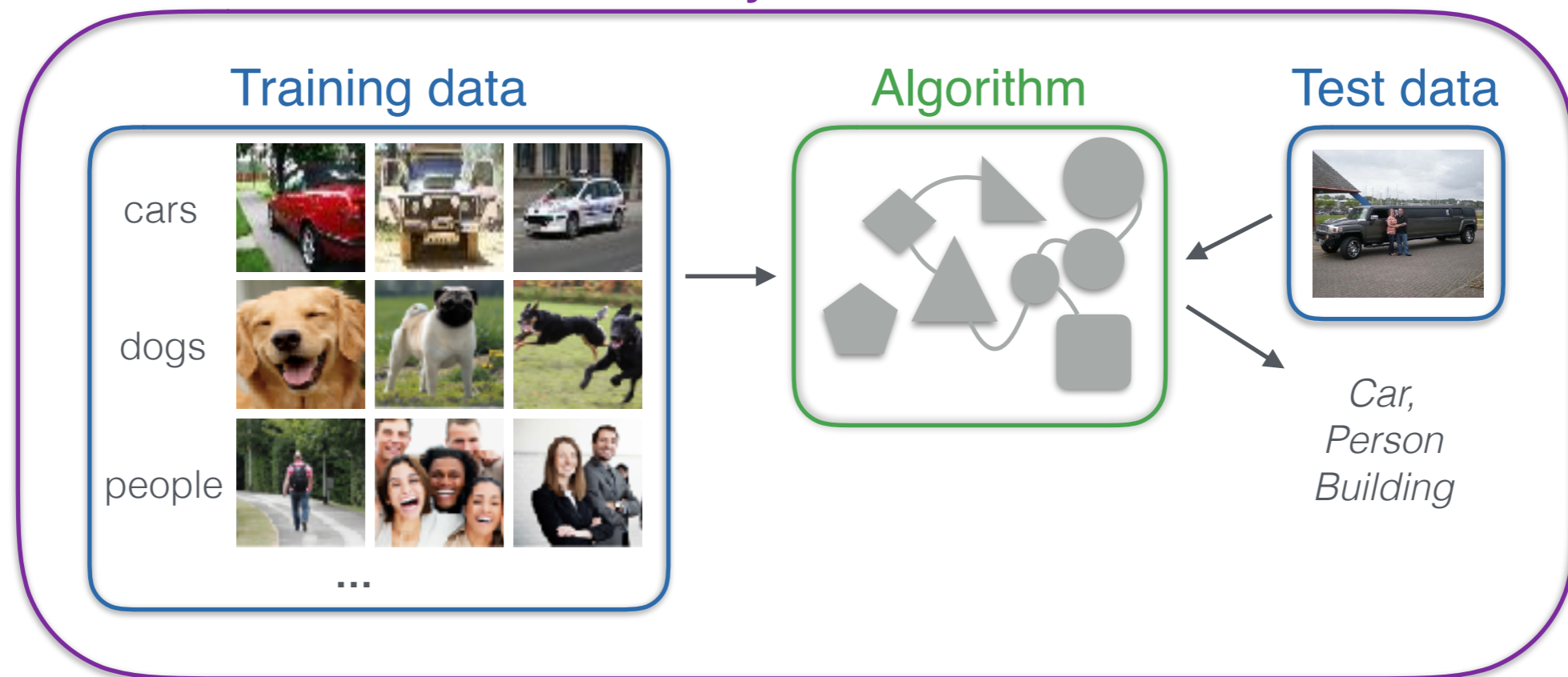


Reading house numbers

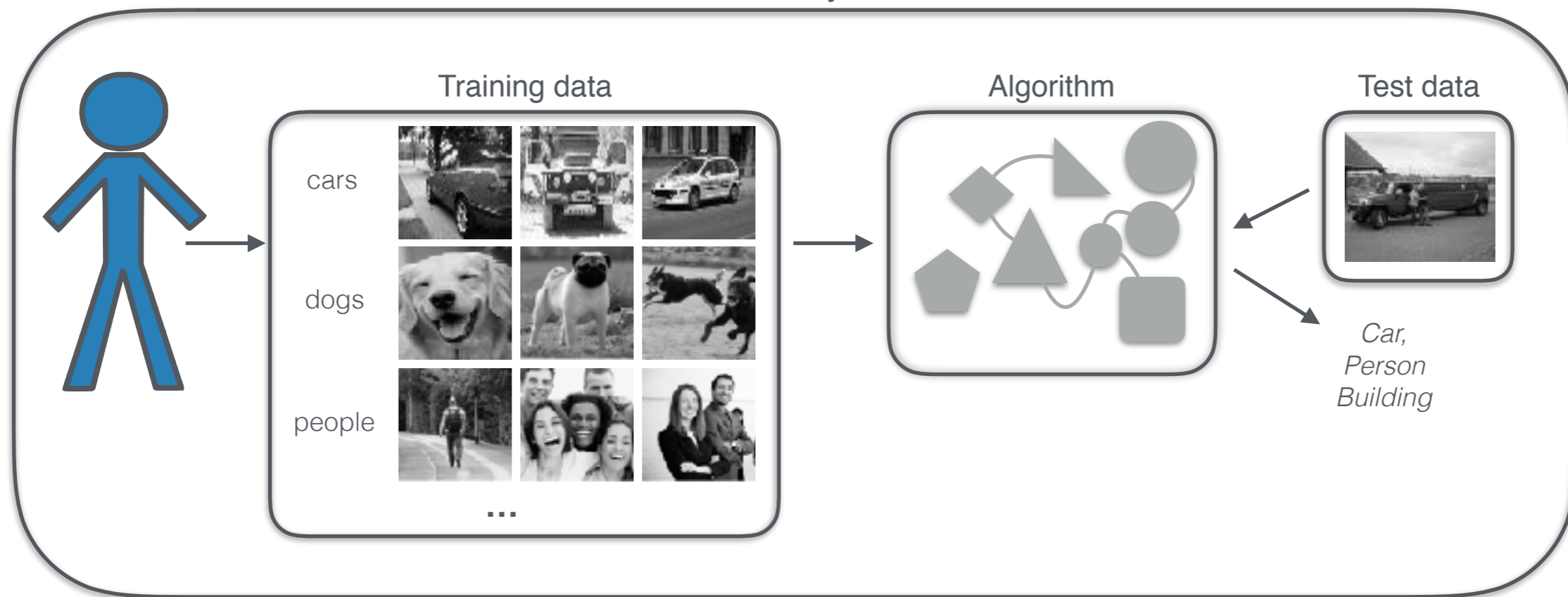


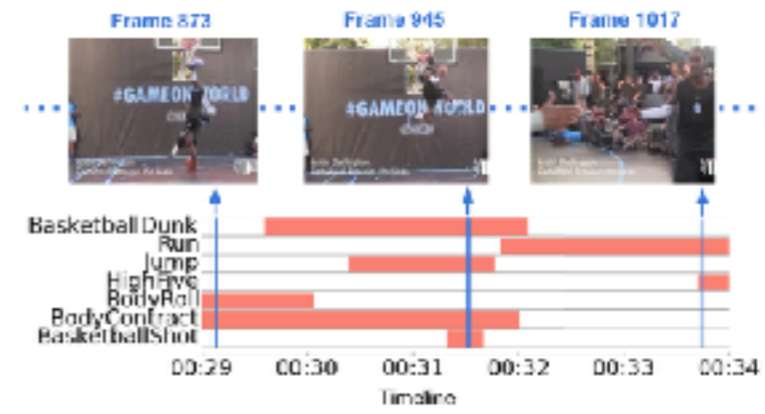
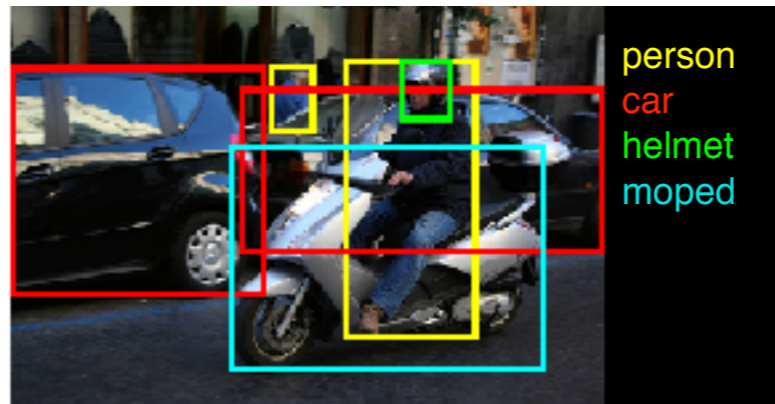
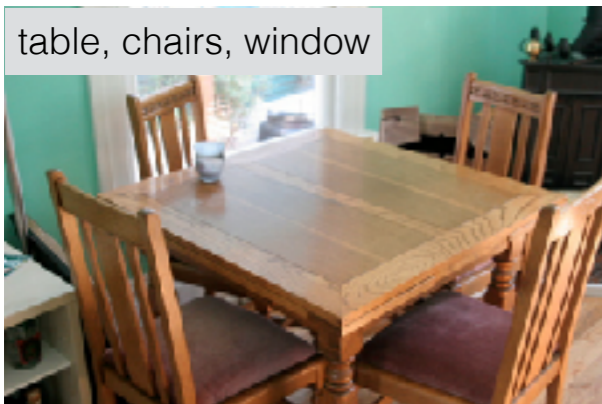
What's missing?

AI vision system

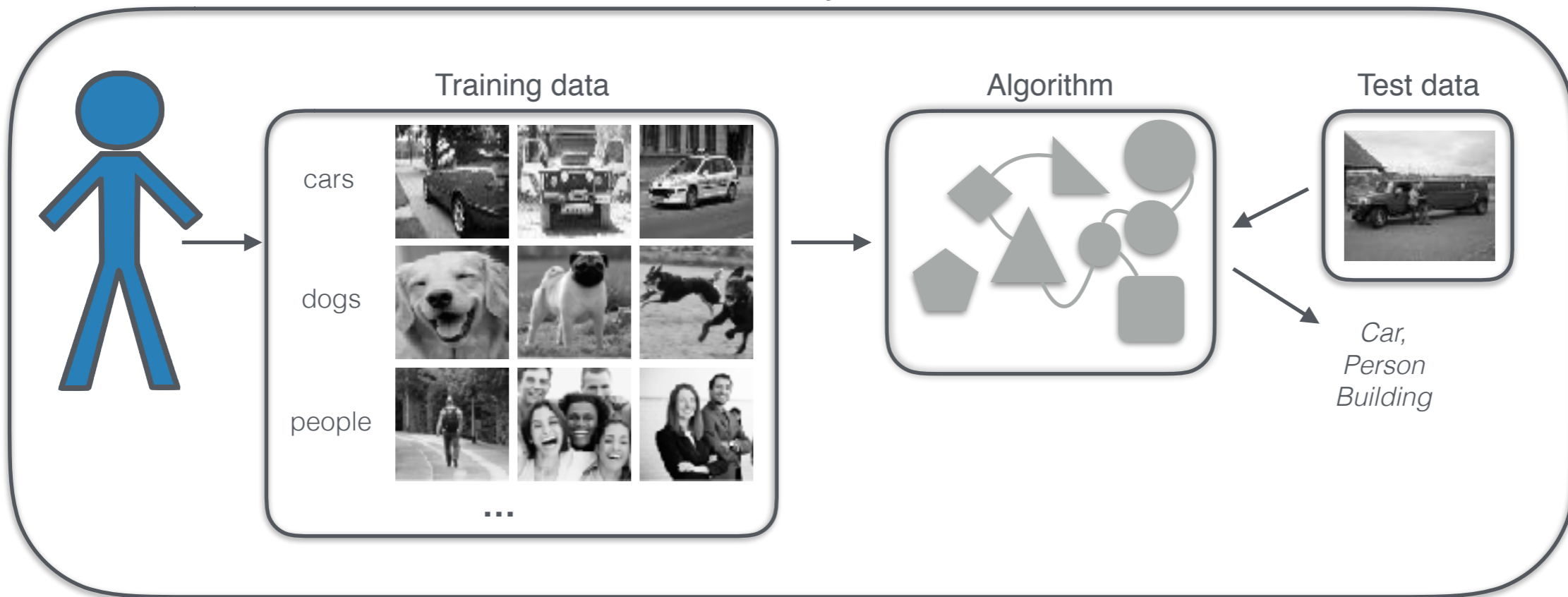


AI vision system






AI vision system



Datasets drive computer vision progress

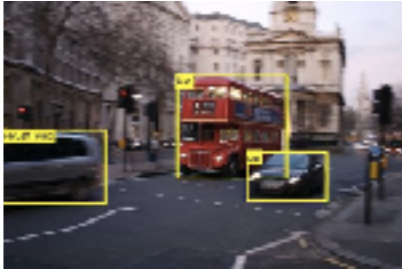
Computer vision capabilities

Caltech 101
[Fei-Fei '04]



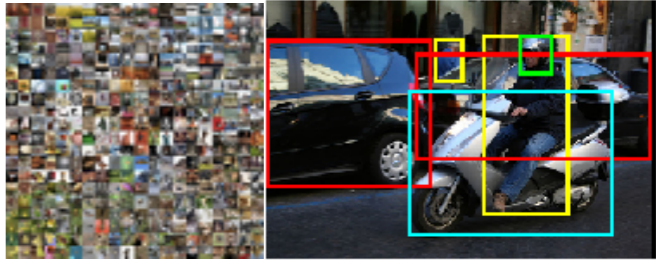
Algorithms:
[Berg '05], [Grauman '05],
[Zhang '06], [Lazebnik '06],
[Jain '08], [Boiman '08],
[Yang '09], [Maji '09]
[Wang '10], [Zhou '10],
[Feng '11], [Jiang '11], ...

PASCAL VOC
[Everingham '07]



Algorithms:
[Chum '07], [Felzenszwalb '08],
[Wang '09], [Harzallah '09],
[Bourdev '09], [Vedaldi '09],
[Lin '09], [Lampert '09],
[Carreira '10], [Wang '10],
[Song '11], [vanDeSande '11], ...



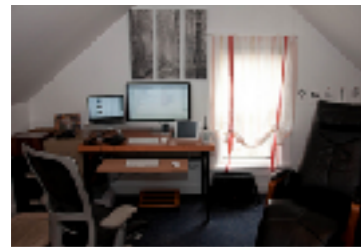

ImageNet
[Deng '09]





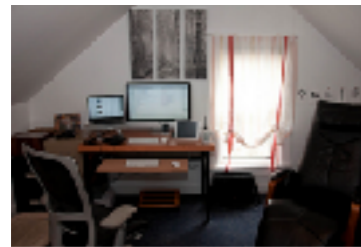

Algorithms:
[Deng '10], [Sanchez '11], [Lin '11],
[Krizhevsky '12], [Zeiler '13], [Wang '13],
[Sermanet '13], [Simonyan '14], [Lin '14],
[Girshick '14], [Szegedy '14], [He '15], ...

Dataset scale and complexity

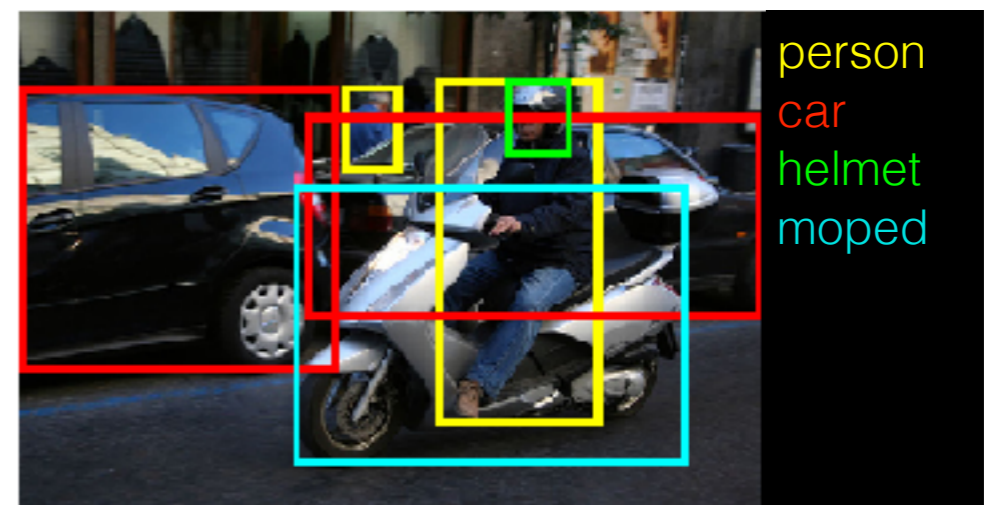
Multi-label annotation

Labels Input	Table	Chair	Bowl
	+	+	-
	+	-	+
	+	+	-
	-	-	-

Multi-label annotation

Labels Input	Table	Chair	Bowl
	+	+	-
	+	-	+
	+	+	-
	-	-	-



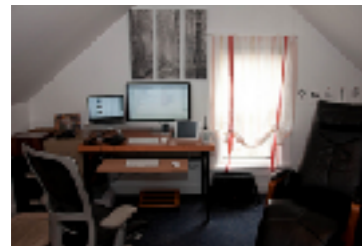

Object detection benchmarks



Alternatively:

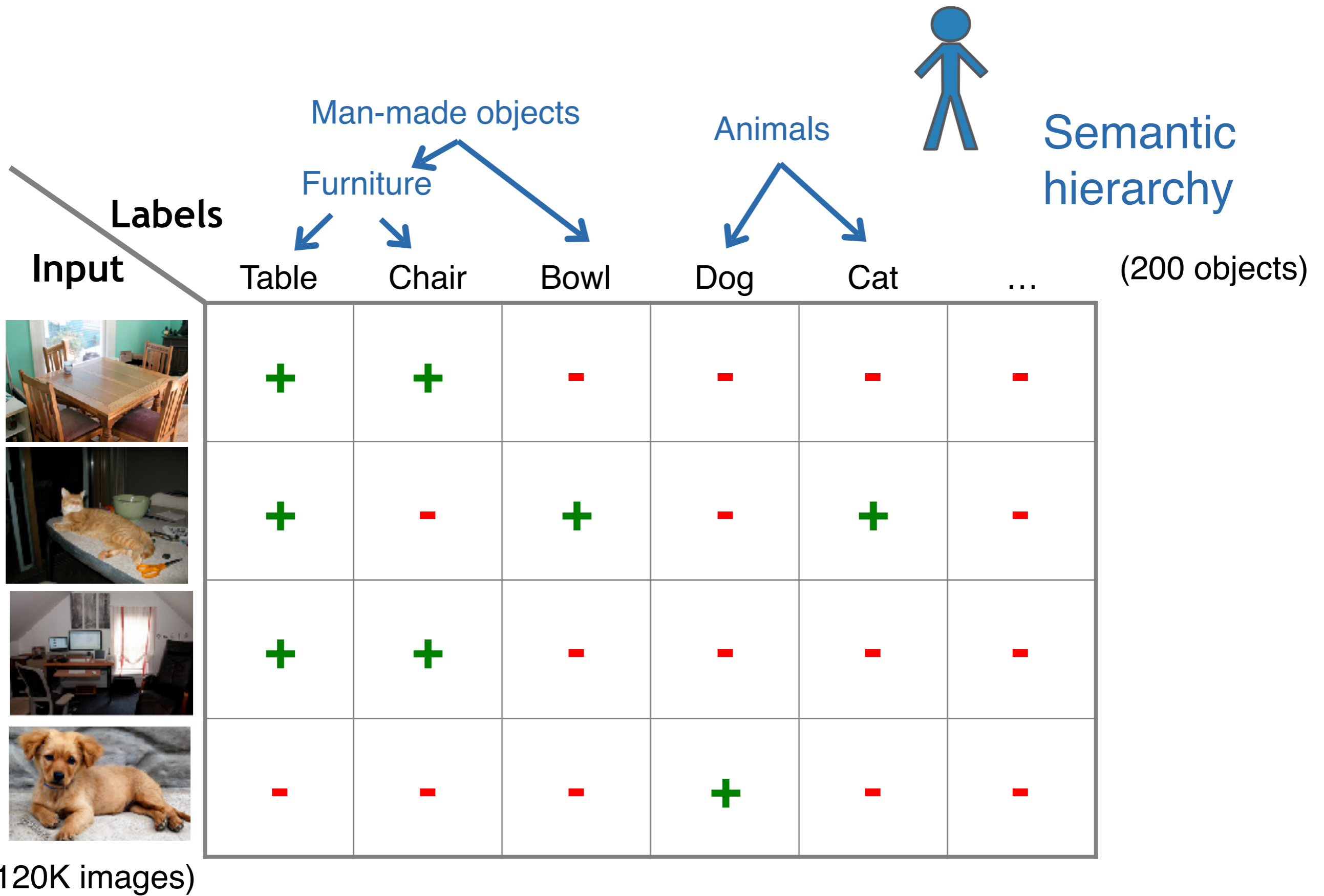
- musical attributes of songs,
- actions in movies,
- sentiments in documents, ...

Multi-label annotation

Labels							(200 objects)
Input	Table	Chair	Bowl	Dog	Cat	...	
	+	+	-	-	-	-	
	+	-	+	-	+	-	
	+	+	-	-	-	-	
	-	-	-	+	-	-	

24 million questions

(120K images)

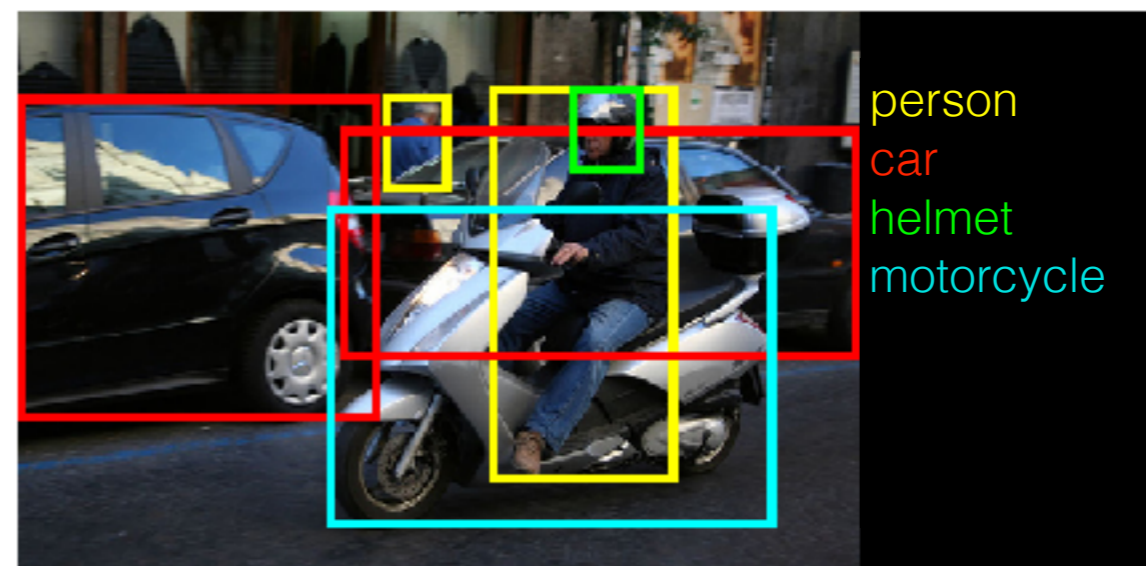
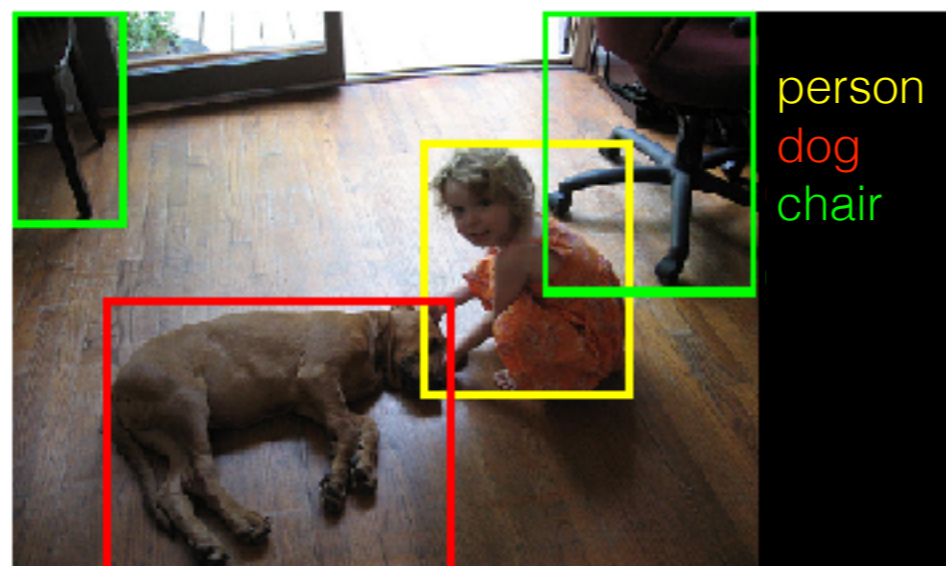


Results:

6.2x savings in multi-label annotation cost

Results:

6.2x savings in multi-label annotation cost
ImageNet **object detection** benchmark



ImageNet large-scale challenge

120,931 images **200** object classes

Compare to PASCAL VOC [EveVanWilWinZis '12]

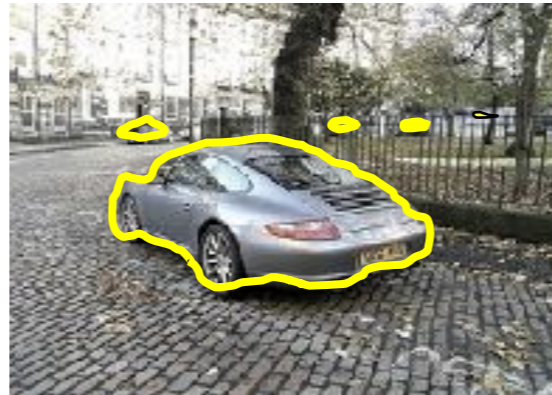
22,591 images **20** object classes

[Deng et al. Scalable Multi-label annotation. CHI'14]

[Russakovsky et al. ImageNet Large Scale Visual Recognition Challenge. IJCV'15]

Working within budget constraints

Task: semantic segmentation



Training annotation options

\$



\$\$\$

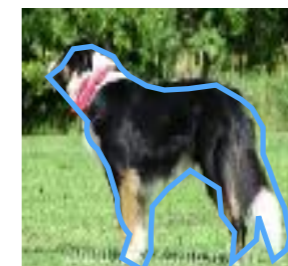
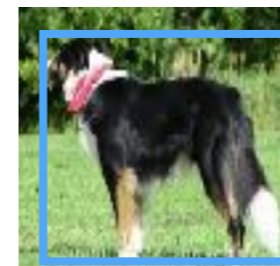
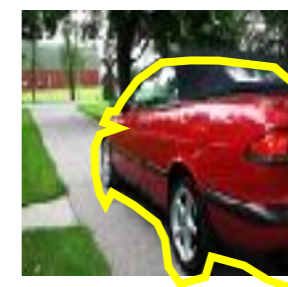
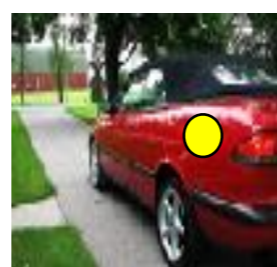
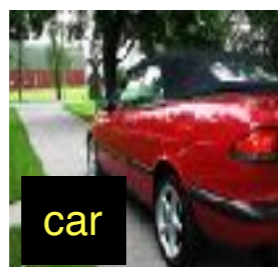


Image-level

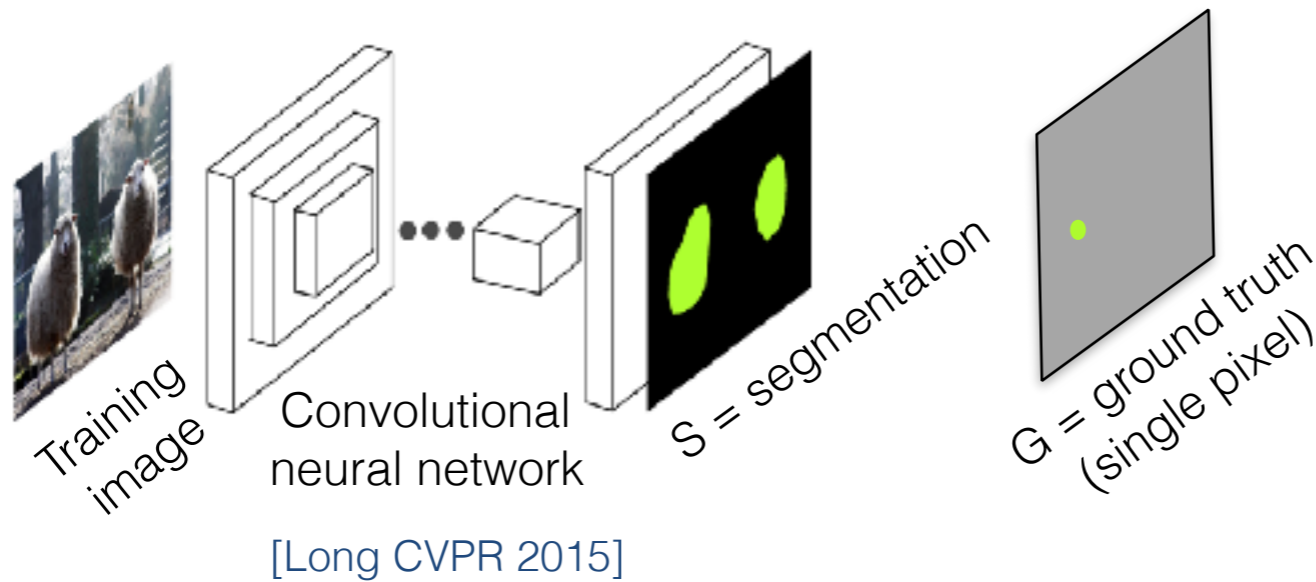
Point-level

Box-level

Full supervision

Working within budget constraints

Training with weak supervision



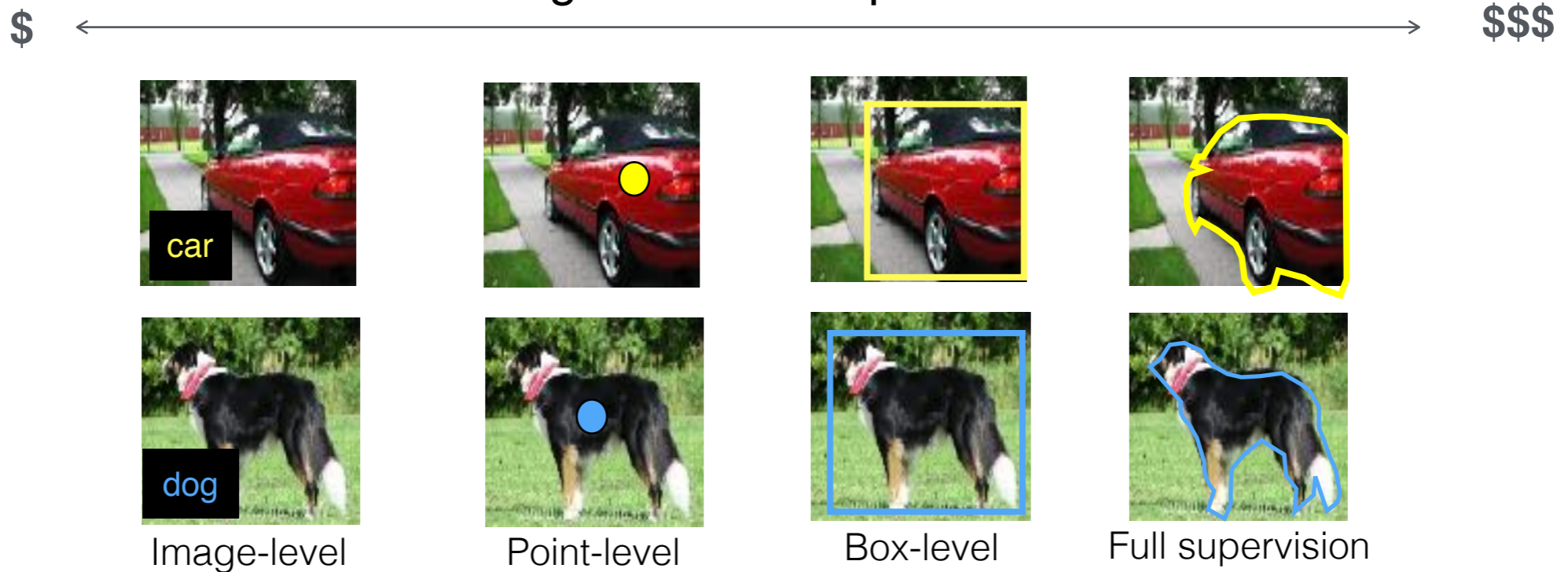
Training loss:

$$\mathcal{L}_{point}(S, G) = - \sum_{i \in \mathcal{G}_s} \log(S_i G_i)$$

Supervised pixels

Probability of target class at pixel i

Training annotation options



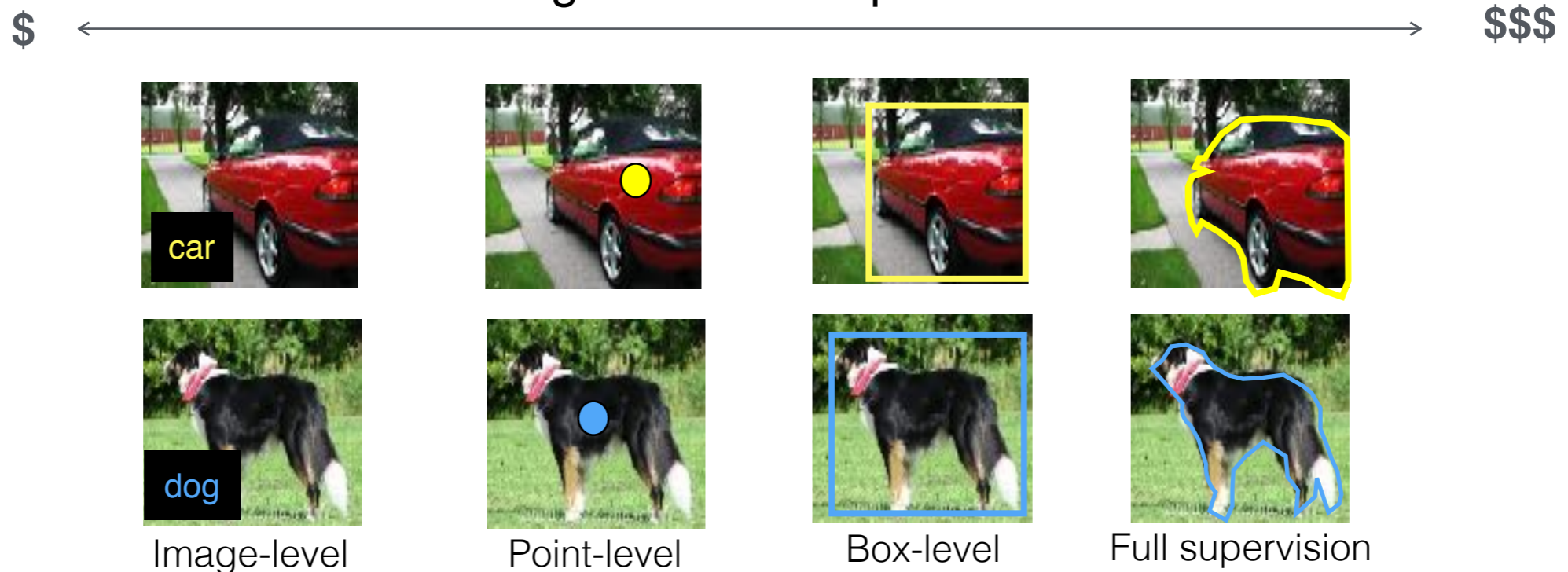
Working within budget constraints

Key result

(on PASCAL VOC, with a fixed budget of human annotation time)

Supervision	mIOU (%)
Full (883 imgs)	22.1
Image-level (10,582 imgs)	29.8

Training annotation options



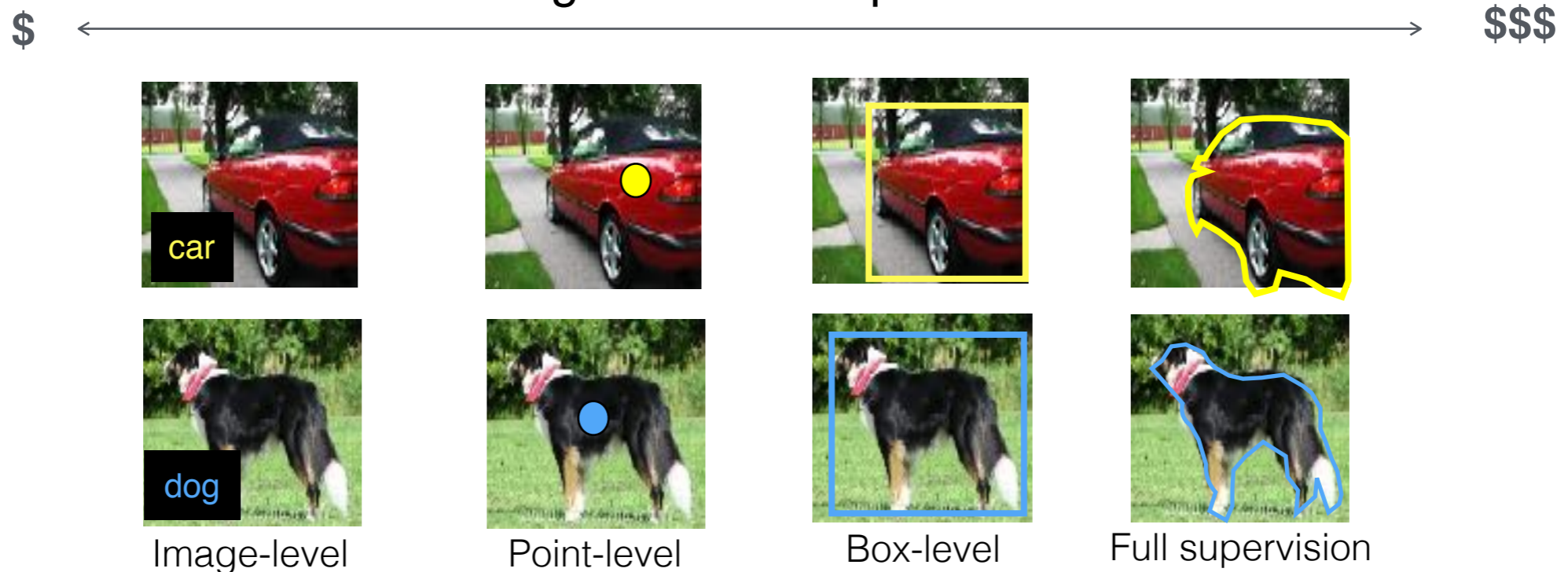
Working within budget constraints

Key result

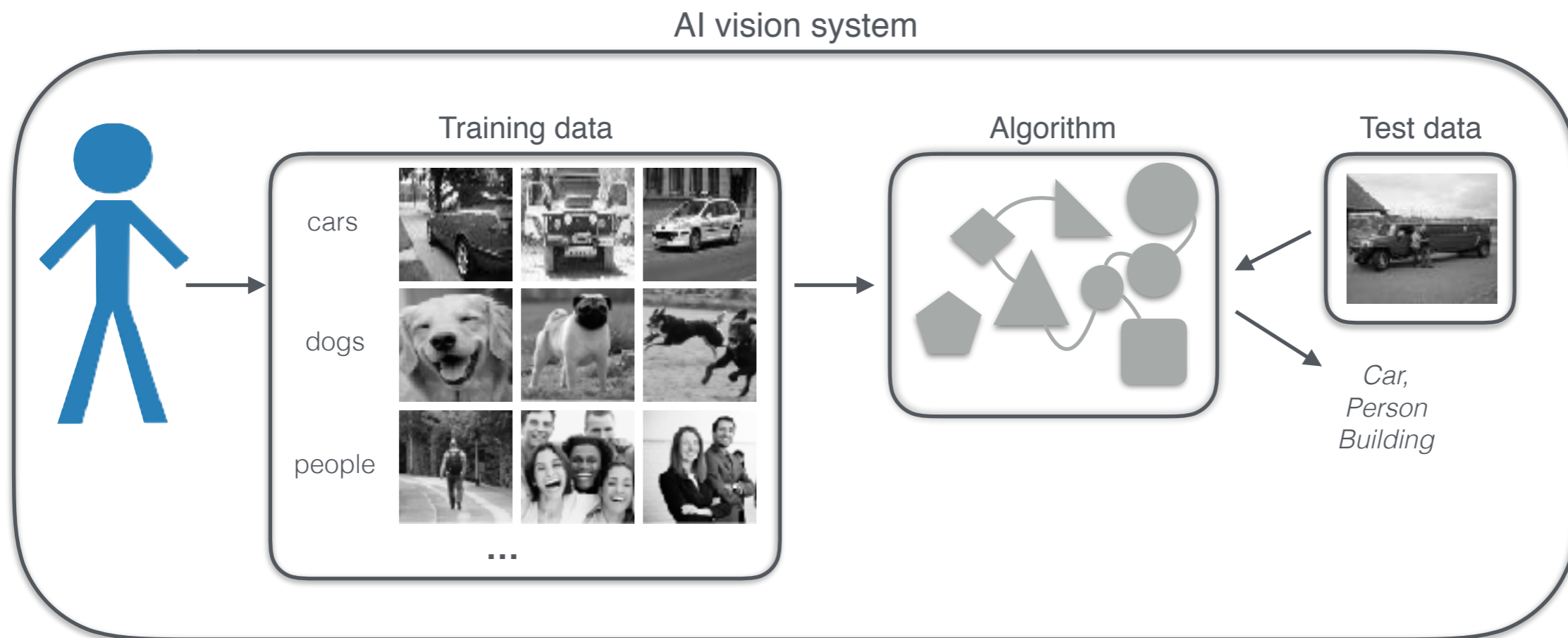
(on PASCAL VOC, with a fixed budget of human annotation time)

Supervision	mIOU (%)
Full (883 imgs)	22.1
Image-level (10,582 imgs)	29.8
Point-level (9,576 imgs)	42.9

Training annotation options



- **Sharing human knowledge with AI**



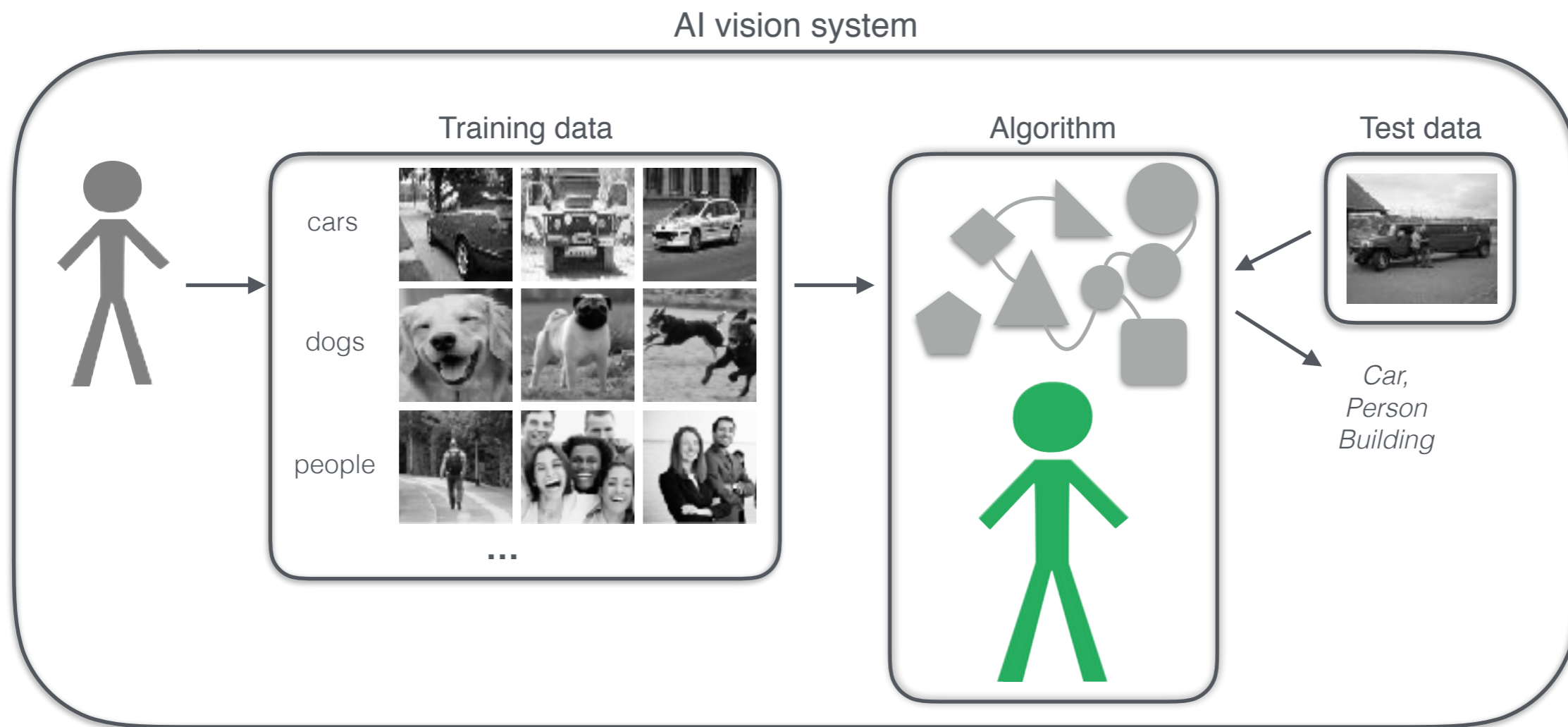
[Bearman et al. What's the point. ECCV'16]

[Sigurdsson et al. Much ado about time. HCOMP'16]

[Russakovsky et al. ImageNet Challenge. IJCV'15]

[Deng et al. Scalable multi-label annotation. CHI'14]

- Sharing human knowledge with AI



[Bearman et al. What's the point. ECCV'16]

[Sigurdsson et al. Much ado about time. HCOMP'16]

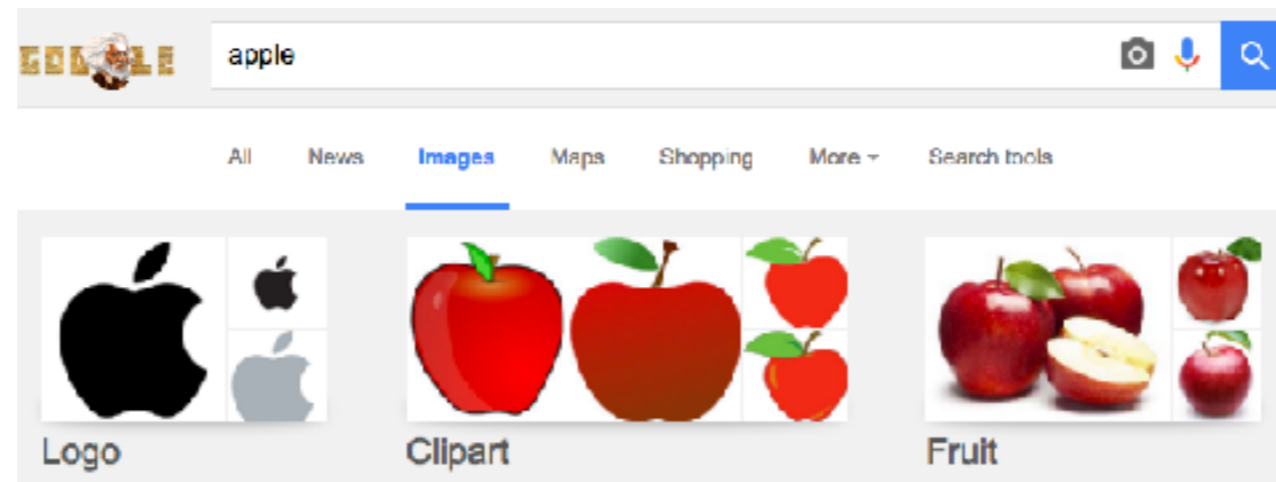
[Russakovsky et al. ImageNet Challenge. IJCV'15]

[Deng et al. Scalable multi-label annotation. CHI'14]

Why human-in-the-loop AI?

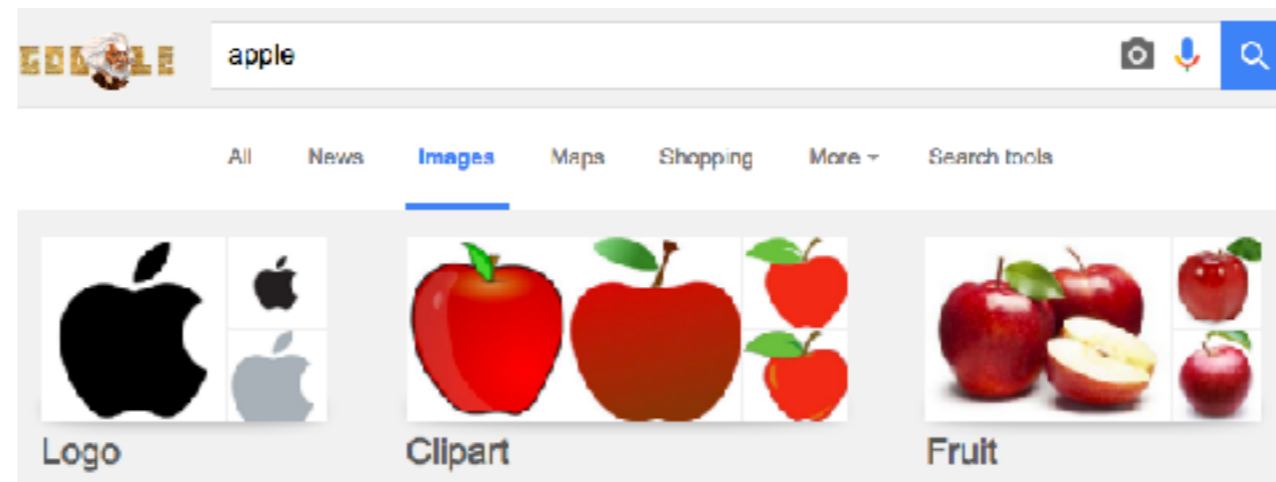
Why human-in-the-loop AI?

Personalization
for unique users



Why human-in-the-loop AI?

Personalization
for unique users

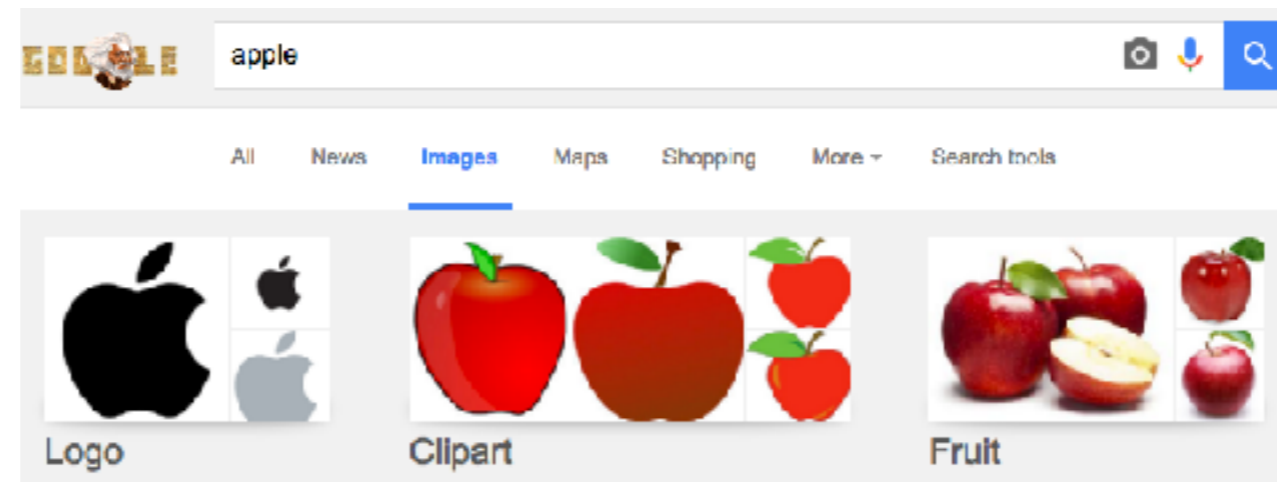


Datasets can't
foresee all
scenarios



Why human-in-the-loop AI?

Personalization
for unique users



Datasets can't
foresee all
scenarios



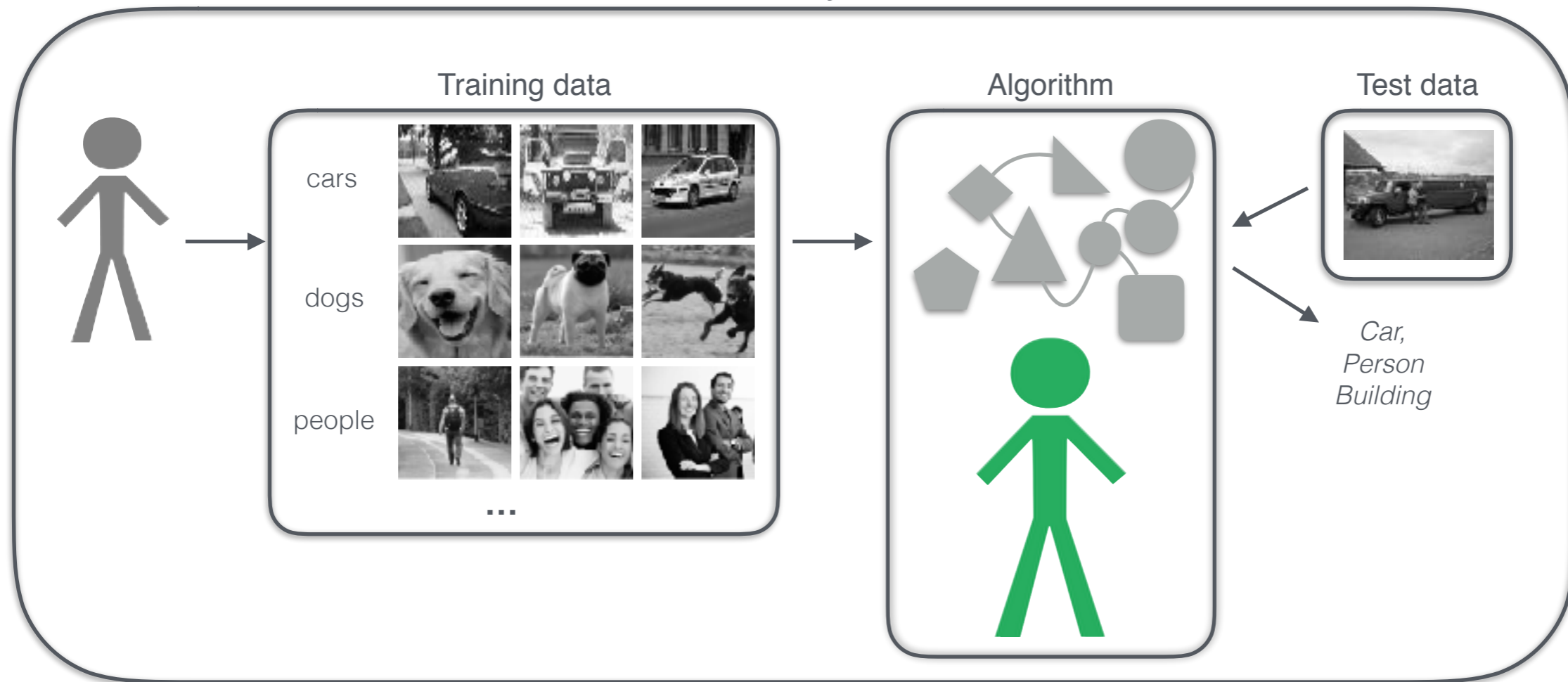
Algorithms are
inherently
imperfect



- Sharing human knowledge with AI

- **Building collaborative and interpretable AI**

AI vision system



[Bearman et al. What's the point. ECCV'16]

[Sigurdsson et al. Much ado about time. HCOMP'16]

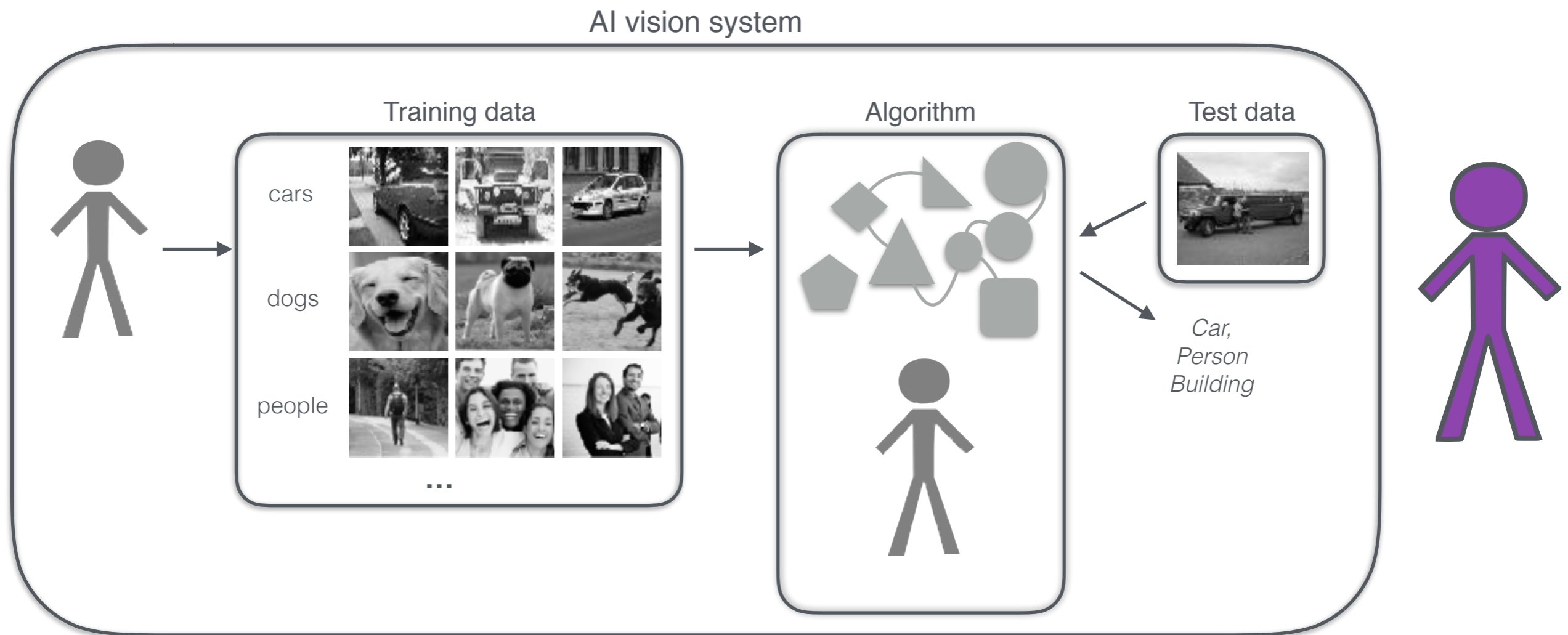
[Russakovsky et al. ImageNet Challenge. IJCV'15]

[Deng et al. Scalable multi-label annotation. CHI'14]

[Russakovsky et al. Best of both worlds. CVPR'15]

- Sharing human knowledge with AI

- Building collaborative and interpretable AI



[Bearman et al. What's the point. ECCV'16]

[Sigurdsson et al. Much ado about time. HCOMP'16]

[Russakovsky et al. ImageNet Challenge. IJCV'15]

[Deng et al. Scalable multi-label annotation. CHI'14]

[Russakovsky et al. Best of both worlds. CVPR'15]

Who is building AI systems?

Who is building AI systems?

- Who is deciding where the data is coming from?

Who is building AI systems?

- Who is deciding where the data is coming from?
- Who is deciding what problems to work on?

Who is building AI systems?

- Who is deciding where the data is coming from?
- Who is deciding what problems to work on?
- Who is deciding what solutions to explore?

Who is building AI systems?

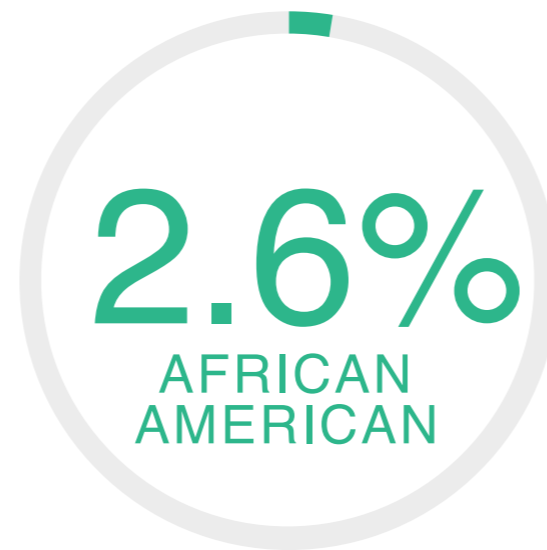
- Who is deciding where the data is coming from?
- Who is deciding what problems to work on?
- Who is deciding what solutions to explore?

**AI will change the world.
Who will change AI?**

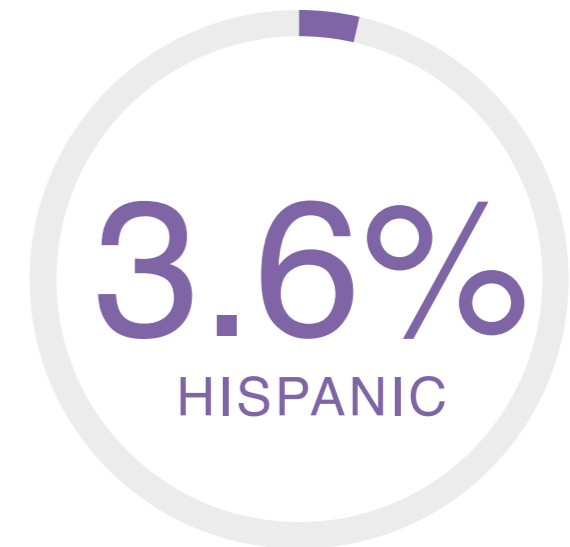
Diversity crisis



Engineering Jobs



Tenure-Track
Engineering Faculty



Tenure-Track
Engineering Faculty

NSF Science and Engineering Indicators, 2014, Association of Engineering Education, 2013

What makes AI special?

- **Rapidly growing revolutionary application-focused field:** critical to avoid bias to particular demographics, needs and values

What makes AI special?

- **Rapidly growing revolutionary application-focused field:** critical to avoid bias to particular demographics, needs and values
- **Fear:** “AI is potentially more dangerous than nukes” [Elon Musk]
“Development of full AI could spell the end of the human race” [Stephen Hawkins]

What makes AI special?

- **Rapidly growing revolutionary application-focused field:** critical to avoid bias to particular demographics, needs and values
- **Fear:** “AI is potentially more dangerous than nukes” [Elon Musk]
“Development of full AI could spell the end of the human race” [Stephen Hawkins]
- **Even more of a diversity crisis:** At Stanford, women account for 32% of computer science majors but only 15% of the AI specialization.

AI4ALL's mission: increase diversity and inclusion in AI education, research, development and policy



“Until this program, I never thought that people who look like me could succeed in computer science and AI.”

- AI4ALL 2016 student



AI4ALL model

- (1) **Educating** the next generation of leaders through summer programs in partnership with universities
- (2) **Supporting** the summer program alumni through ongoing education, mentorship and career opportunities
- (3) **Expanding** awareness by funding alumni ambassadors to educate the broader community



AI4ALL team



Tess Posner
Executive
Director



Nicole Halmni
Project+Marketing
Manager



Wells Santo
Education Manager



Fei-Fei Li
Co-founder,
board member



Olga Russakovsky
Co-founder,
board member



Rick Sommer
Co-founder,
board member



Rab Govil
Board member



AI4ALL seed funders



Melinda Gates
Pivotal Ventures



Jensen and Lori Huang
Jensen + Lori Huang
Foundation



Initial results: Stanford program 2015-2017



- **83.3%** of students are proud of what they build during the program
- **75%** of students feel confident that they can undertake a similar AI project in the future
- After the program, students are **20%** more likely to feel like they are part of a community in computer science



AI4ALL 2018 summer programs

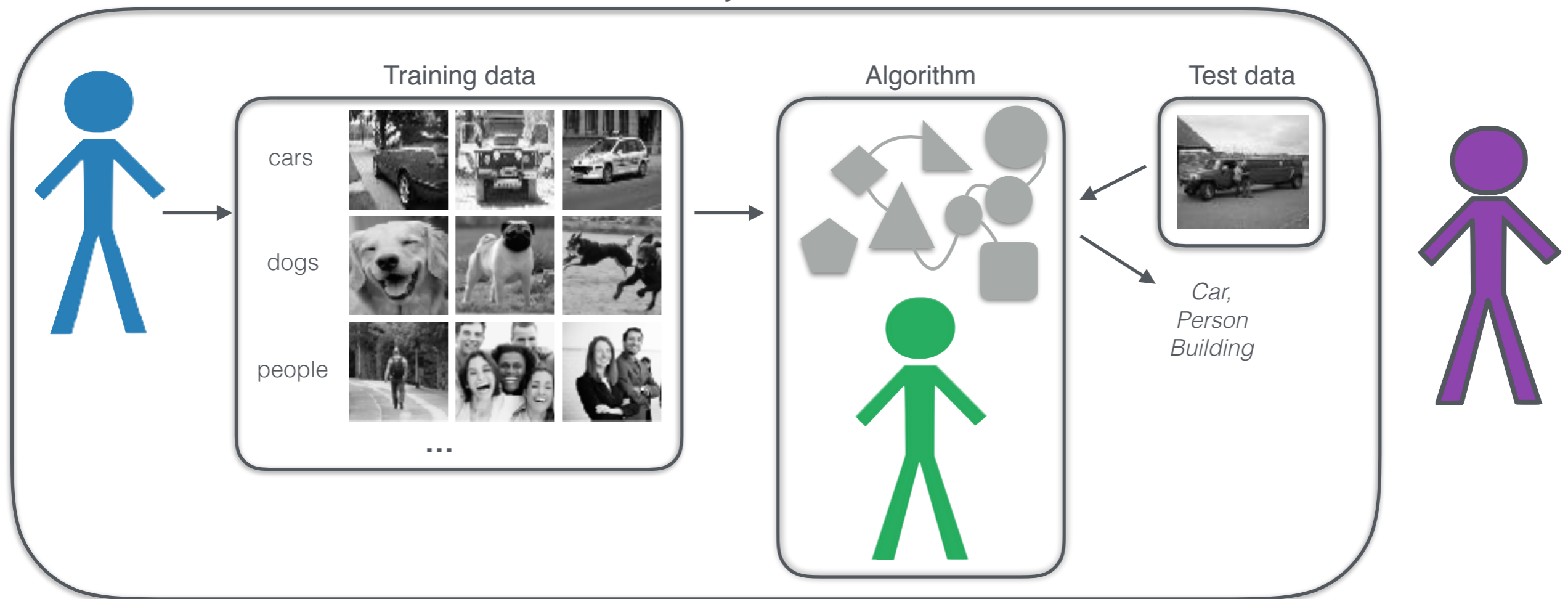


- **Sharing human knowledge with AI**

- **Building collaborative and interpretable AI**

- **Creating diverse and inclusive AI**

AI vision system



[Bearman et al. What's the point. ECCV'16]

[Sigurdsson et al. Much ado about time. HCOMP'16]

[Russakovsky et al. ImageNet Challenge. IJCV'15]

[Deng et al. Scalable multi-label annotation. CHI'14]

[Russakovsky et al. Best of both worlds. CVPR'15]

[Vachovsky et al. Towards diversity. SIGCSE'16]

<http://ai-4-all.org>

