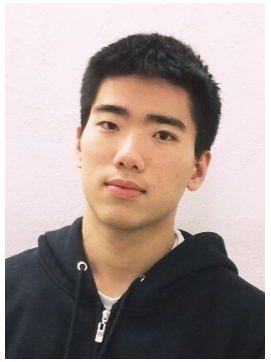
DecompoVision: Reliability Analysis of Machine Vision Components through Decomposition and Reuse

ESEC/FSE 2023

Boyue Caroline Hu

Lina Marsso

Nikita Dvornik



Huakun Shen

Marsha Chechik

UNIVERSITY OF TORONTO

Motivation

Machine Learning models solve a variety of vision tasks in many domains

- Atomic vision task: one type of output
  - e.g., image classification -> class label
- Complex vision task: multiple types of outputs
  - e.g., object detection -> bounding box and class label
  - e.g., instance segmentation -> bounding box, class label and segmentation masks

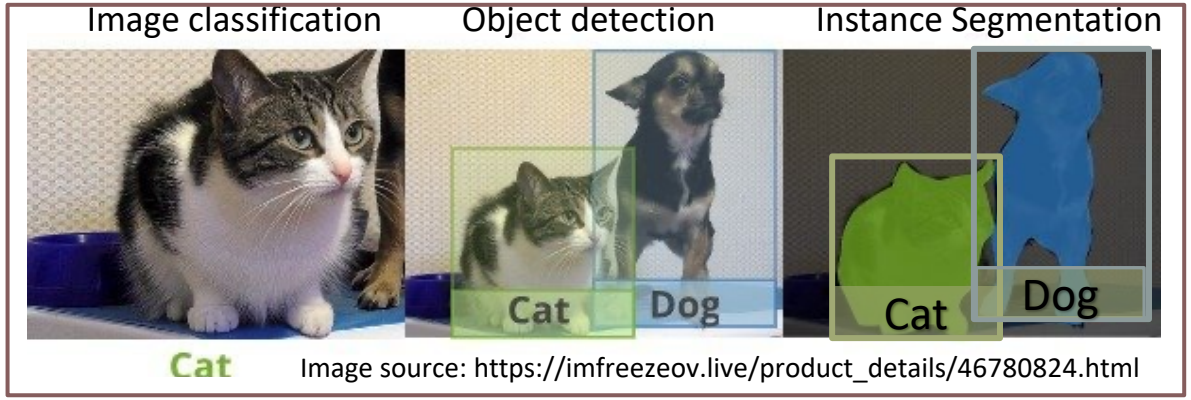


Image source: https://imfreezeov.live/product_details/46780824.html

ML reliability: stability facing perturbations in the input

- Performance comparison with humans
- Reliability requirements
- Testing
- …

Reliability of atomic tasks ✓

Reliability of complex vision tasks?

DecompoVision: a modular reliability framework through **decomposing analysis** of vision tasks

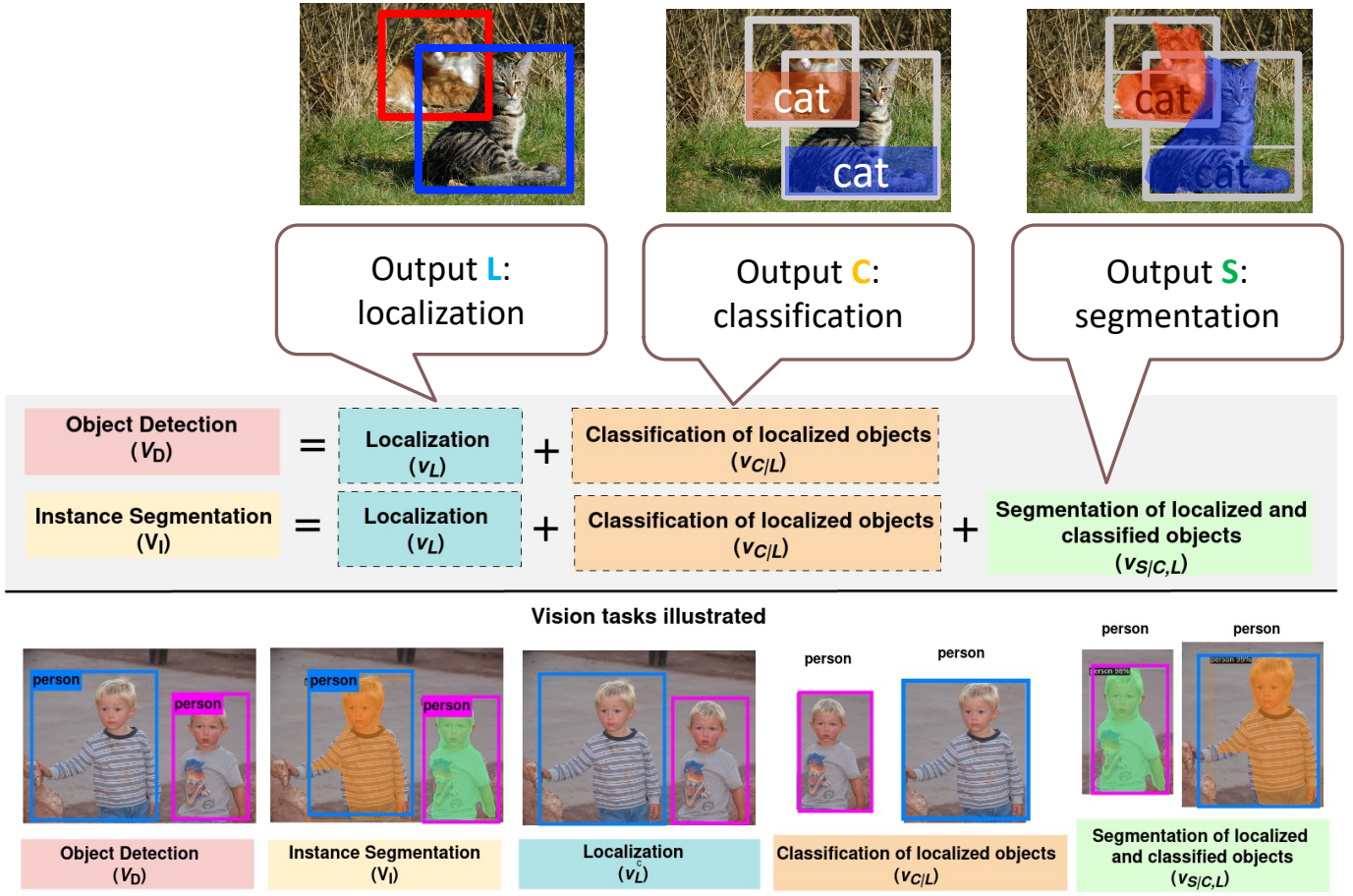
Vision Task Decomposition

"Many complex vision tasks can be represented as a *sequence* of atomic vision subtasks" [Haralick, 1992]

Decomposition Principle (DP): decompose the **problem of solving a complex vision task** into solving the corresponding atomic subtasks of each output type

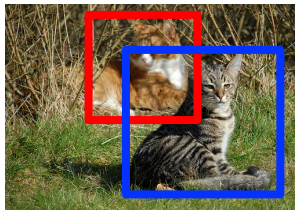
Other decomposable complex vision tasks:

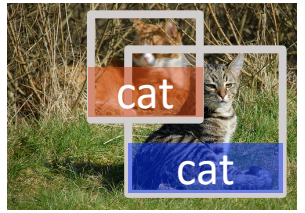
- human-object interaction detection
- trajectory prediction
- language visual grounding
- …

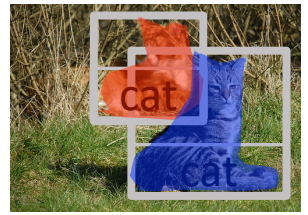


Robert M. Haralick. 1992. Performance Characterization in Computer Vision. In Proceedings of the British Machine Vision Conference, BMVC 1992, Leeds, UK, September, 1992. BMVA Press, 1–8. https://doi.org/10.5244/C.6.1

Vision Task Decomposition







Output **L**: localization

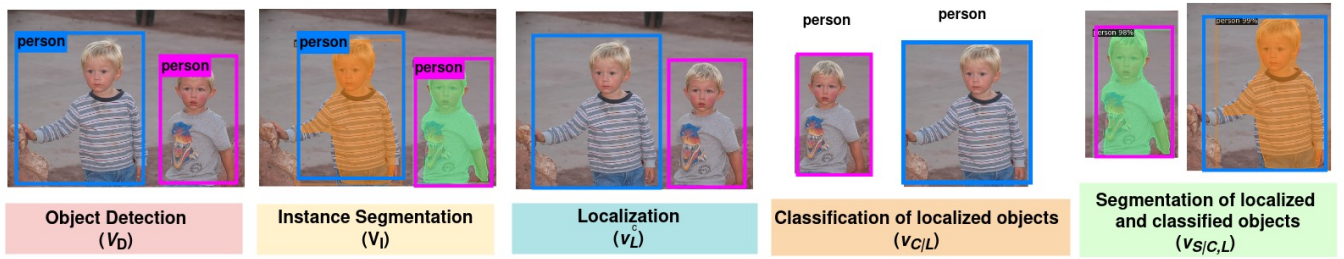
Output **C**: classification

Output **S**: segmentation

Decomposition Principle (DP): decompose the **problem of solving a complex vision task** into solving the corresponding atomic subtasks of each output type

Object Detection (V_D)	=	Localization (v_L)	+	Classification of localized objects ($v_{C\|L}$)		
Instance Segmentation (V_I)	=	Localization (v_L)	+	Classification of localized objects ($v_{C\|L}$)	+	Segmentation of localized and classified objects ($v_{S\|C,L}$)

Vision tasks illustrated



Why **C**|**L**?

Conditional probability chain rule: the probability of observing all the outputs can be decomposed into probabilities of observing a sequence of individual outputs

For object detection

$$P(V_D) = P(C, L) = P(L) * P(C \mid L)$$

i.e., $P(V_D) = P(v_L) * P(v_{C|L})$

Modular Analysis through Analysis Decomposition

Decomposition Principle (DP): decompose the **problem of solving a complex vision task** into solving the corresponding atomic subtasks of each output type

Through DP, the **performance analysis** can also be decomposed through performance metrics!

Precision for localization (v_L)

Precision for classification given localization ($v_{C|L}$)

$$Precision_D = Precision_L * Precision_{C|L}$$

Precision for object detection (V_D)

See the proof in our paper

Decomposing performance analysis allows decomposition of the entire analysis workflow,
e.g., comparing with humans, requirements, testing

Such decomposition...

- … allows reusing of analysis artifact
- … allows modular development
- … leads to a deeper insight into ML reliability

Note: model decomposition **NOT** required, can be black-box

Decomposing Comparison with Humans

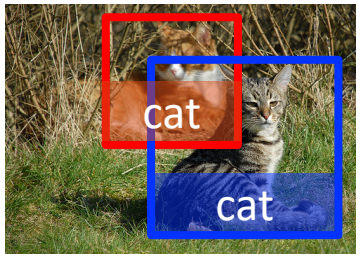
"A trustworthy ML model should at least reach human performance" [Firestone, 2020]

Human performance experiments decomposed with DP:

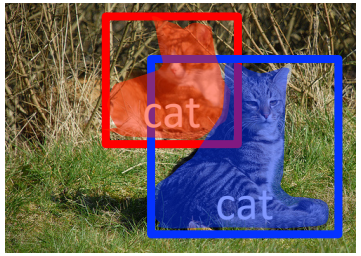
- A fair comparison between human and machine
- Reuse of experiment data

Without decomposition

Object detection (v_D)



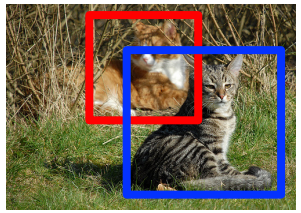
Instance segmentation (v_I)



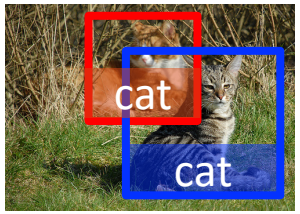
Can reduce the experiment time by 33%

Time limit for each subtask

With decomposition



+

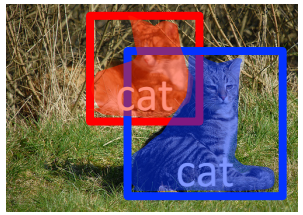


v_L: localizing objects

$v_{C|L}$: classifying localized objects

Reuse results for v_L and $v_{C|L}$

+



$v_{S|C,L}$: draw outline of classified and localized object

Chaz Firestone. 2020. Performance vs. Competence in Human–Machine Comparisons. Proceedings of the National Academy of Sciences 117, 43 (2020), 26562–26571.

Decomposing Reliability Requirements

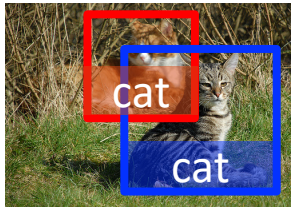
ML reliability requirement REQ(M)*:* performance measured with metric M should not be affected by perturbations [Hu et. al, 2022]

REQ_L : localization performance measured with the decomposed $Precision_L$ should not be affected.

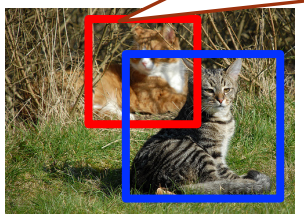
REQ_L and $REQ_{C|L}$ for objection can potentially be reused for instance segmentation.

Without decomposition

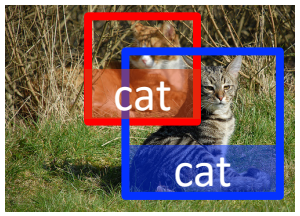
Object detection (V_D)



REQ_D ($Precision_D$)

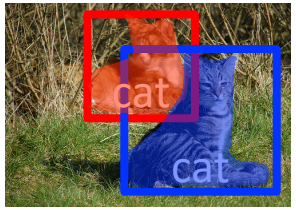


+

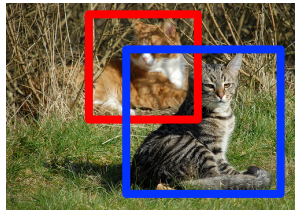


v_L: REQ_L ($Precision_L$) $v_{C|L}$: $REQ_{C|L}$ ($Precision_{C|L}$)

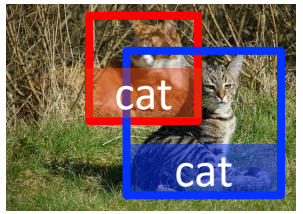
Instance segmentation (V_I)



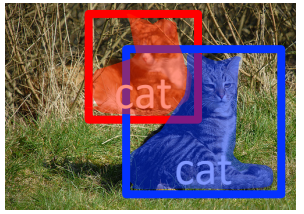
REQ_I ($Precision_I$)



+



+



v_L: REQ_L ($Precision_L$) $v_{C|L}$: $REQ_{C|L}$ ($Precision_{C|L}$) $v_{S|C,L}$: $REQ_{S|C,L}$ ($Precision_{S|C,L}$)

Theorem: satisfying REQ_L AND $REQ_{C|L}$ implies satisfying REQ_D See the proof in our paper

Boyue Caroline Hu, Lina Marsso, Krzysztof Czarnecki, Rick Salay, Huakun Shen, and Marsha Chechik. 2022. If a Human Can See It, So Should Your System: Reliability Requirements for Machine Vision Components. In Proceedings of the 44th International Conference on Software Engineering (ICSE'2022), Pittsburgh, USA. ACM.

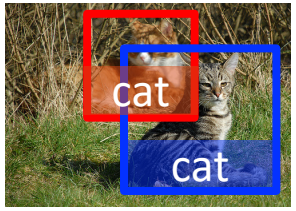
Decomposing Reliability Testing

Enabled from decomposing

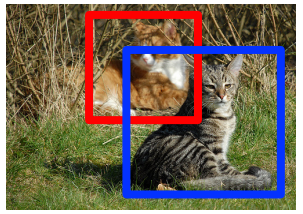
- performance metric
- reliability requirements

Without decomposition | **With decomposition**

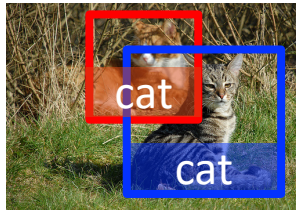
Object detection (V_D)



$Precision_D$ or REQ_D



+

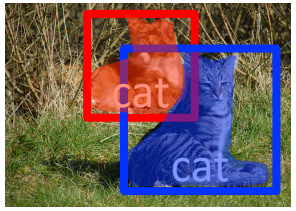


V_L: $Precision_L$ or REQ_L $V_{C|L}$: $Precision_{C|L}$ or $REQ_{C|L}$

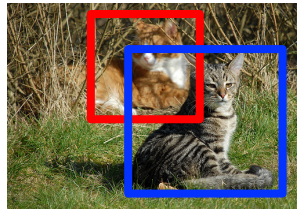
Little computational overhead

➡ $Precision_D$ or REQ_D

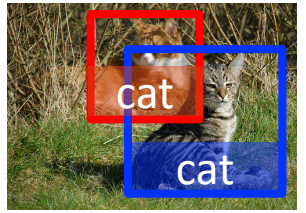
Instance segmentation (V_I)



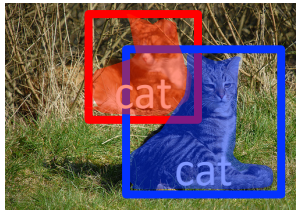
$Precision_I$ or REQ_I



+



+



V_L: $Precision_L$ or REQ_L $V_{C|L}$: $Precision_{C|L}$ or $REQ_{C|L}$ $V_{S|C,L}$: $Precision_{S|C,L}$ or $REQ_{S|C,L}$

More insights on model performances on subtasks

➡ $Precision_I$ or REQ_I

Decomposing Reliability Testing

Detecting people with the PASCAL-VOC dataset:

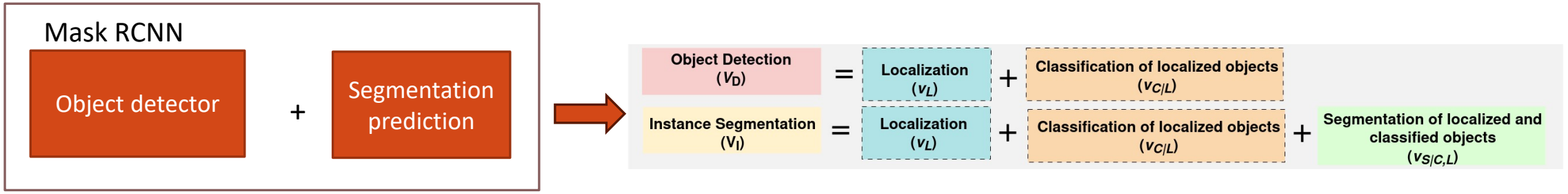
Model name	AP for V_D	Satisfaction of REQ_L	Satisfaction of $REQ_{C|L}$
R101-C4-3x	**0.39**	**0.77**	**0.84**
R101-DC5-3x	**0.39**	**0.81**	**0.73**

Metric for measuring degree of satisfying a requirement. [0, 1]

- Same Average Precision (AP) value for object detection
- Different ability for localizing objects and classifying localized objects
- Subtask analysis allows:
  - improvements of model performance
  - model selection for different application

Decomposing Reliability Testing

Multi-stage ML models, such as Mask RCNNs, correspond to DP



Thus, analysis artifacts (tests, metrics, requirements) of the object detector can be reused for instance segmentation.

	With Decomposition							No Decomposition	
	Object Detection			Instance Segmentation				O. D.	I. S.
Runtime (s)	507.49	2.33	1.00	144.35	0.63	34.97	0.62	511.26	692.21
Peak Mem (MB)	61.36	61.36	61.36	61.36	61.36	61.36	61.36	61.36	61.36

Limitations and Future Work

Limitations

Applicable to only decomposable complex vision tasks and metrics

Satisfying all decomposed requirements
=> satisfying the requirement for the complex vision task
The other direction might not hold

Future Work

Decomposition for a broader range of vision tasks

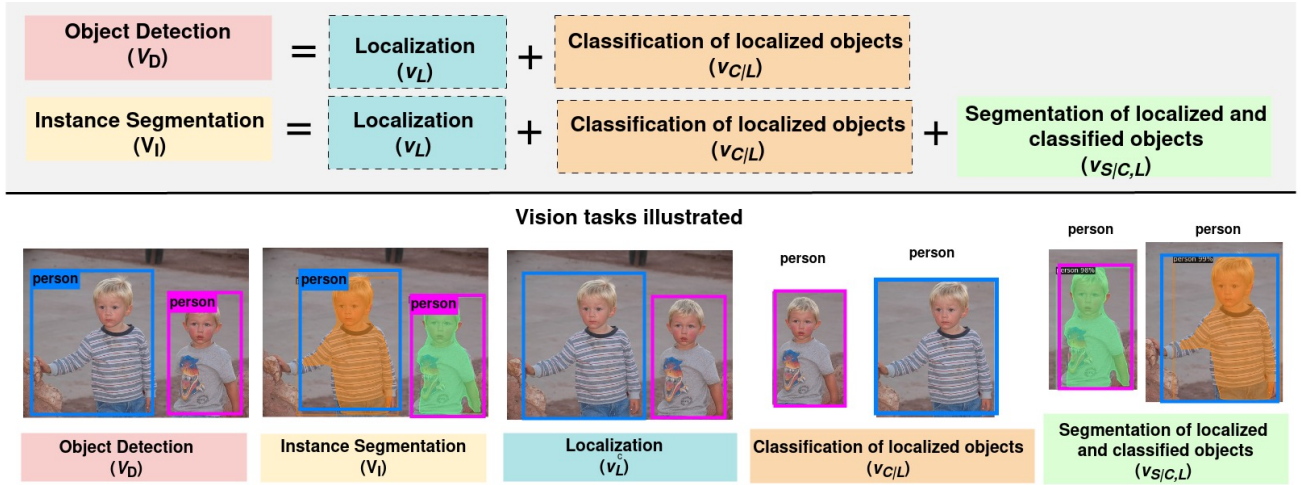
- Improve satisfaction measures
- Improve requirements decomposition

Use of decomposition for other reliability analyses, e.g., verification

Conclusion

DecompoVision: Reliability Analysis of Machine Vision Components through Decomposition and Reuse

- **Decomposition Principle (DP)**: complex vision tasks can be decomposed into a sequence of atomic subtasks
- DecompoVision – a modular framework that **decomposes reliability analysis** of ML performing complex vision tasks
- DecompoVision's modularity enables:
  - reuse analysis artifacts
  - getting deeper insights about the subtask reliability



UNIVERSITY OF TORONTO

caroline + 6 e's

Github:
https://github.com/carolineeeeeee/DecompoVision

I'm on the job market!
Contact me @
boyue@cs.toronto.edu

Boyue Caroline Hu

Thank you!