How to navigate through the ML research information flood

Dmytro Mishkin. Faculty of Electrical Engineering, CTU in Prague

CFNet: Learning Correlation Functions for One-Stage Panoptic Segmentation

Yifeng Chen, Wenqing Chu, Fangfang Wang, Ying Tai, Ran Yi, Zhenye Gan, Liang Yao, Chengjie Wa...

Recently, there is growing attention on one-stage panoptic segmentation methods which aim to segment instances and stuff jointly within a fully convolutional pipeline efficiently. However, most of the existing works directly feed the backbone features to various segmentation heads ignoring the demands for semantic and instance segmentation are different: The former needs semantic-level discriminative features, while the latter requires features to be distinguishable across instances. To alleviate this, we propose to first predict semantic-level and instance-level correlations among different locations that are utilized to enhance the backbone features, and then feed the improved discriminative features into the corresponding segmentation heads,

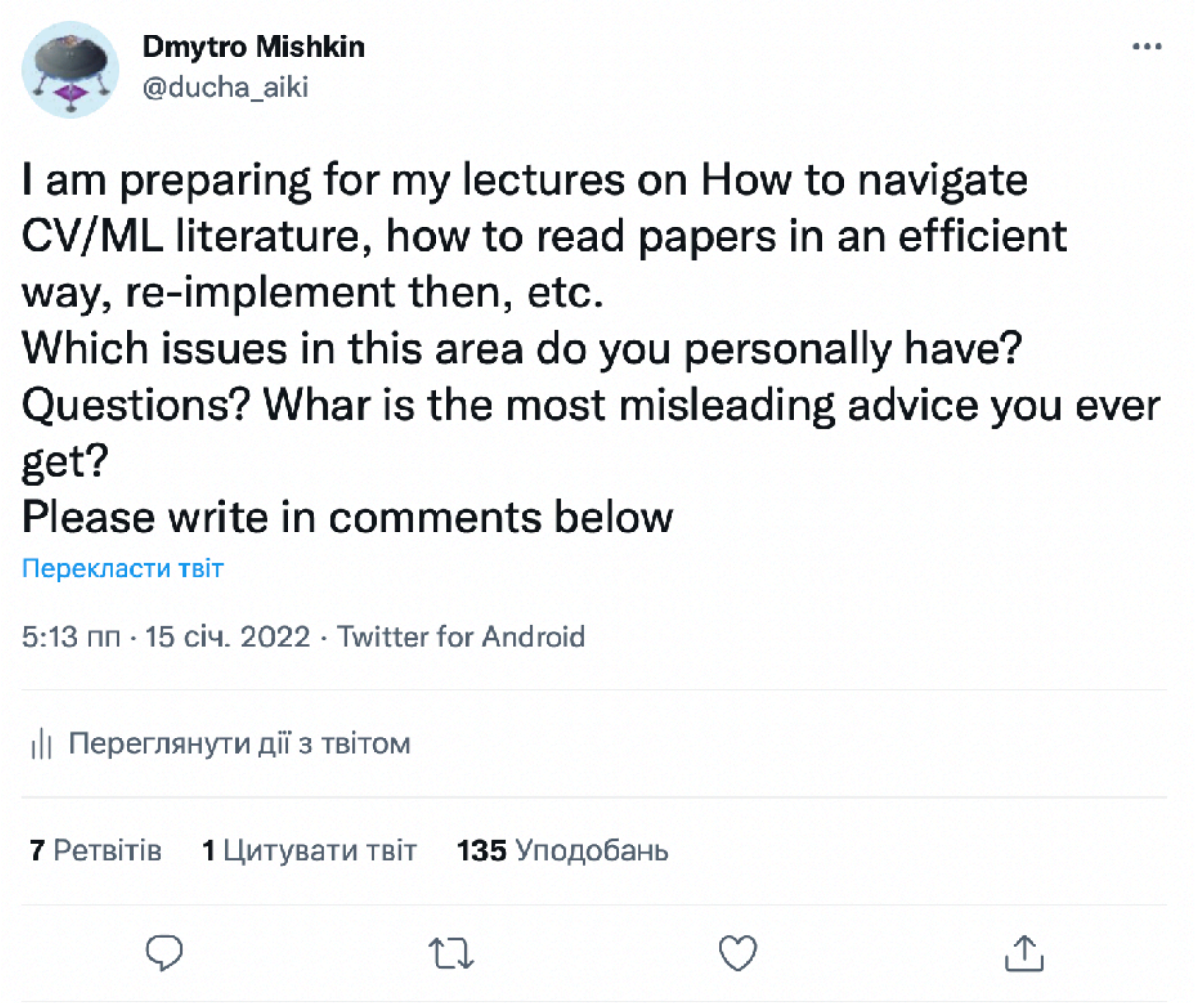
CFNet: Learning Correlation Functions for One-Stage Panoptic Segmentation

Yifeng Chen, Wenqing Chu, Fangfang Wang, Ying Tai, Ran Yi, Zhenye Gan, Liang Yao, Chengjie Wa...

Recently, there is growing attention on one-stage panoptic segmentation methods which aim to segment instances and stuff jointly within a fully convolutional pipeline efficiently. However, most of the existing works directly feed the backbone features to various segmentation heads ignoring the demands for semantic and instance segmentation are different: The former needs semantic-level discriminative features, while the latter requires features to be distinguishable across instances. To alleviate this, we propose to first predict semantic-level and instance-level correlations among different locations that are utilized to enhance the backbone features, and then feed the improved discriminative features into the corresponding segmentation heads,

Foreword

Foreword



Dmytro Mishkin
@ducha_aiki

I am preparing for my lectures on How to navigate CV/ML literature, how to read papers in an efficient way, re-implement then, etc.
Which issues in this area do you personally have? Questions? Whar is the most misleading advice you ever get?
Please write in comments below

Перекласти твіт

5:13 пп · 15 січ. 2022 · Twitter for Android

Переглянути дії з твітом

7 Ретвітів 1 Цитувати твіт 135 Уподобань

Foreword

Dmytro Mishkin
@ducha_aiki

I am preparing for my lectures on How to navigate CV/ML literature, how to read papers in an efficient way, re-implement then, etc.
Which issues in this area do you personally have? Questions? Whar is the most misleading advice you ever get?
Please write in comments below

Перекласти твіт

5:13 пп · 15 січ. 2022 · Twitter for Android

Переглянути дії з твітом

7 Ретвітів **1** Цитувати твіт **135** Уподобань

Foreword

Dmytro Mishkin
@ducha_aiki

I am preparing for my lectures on How to navigate CV/ML literature, how to read papers in an efficient way, re-implement then, etc.
Which issues in this area do you personally have?
Questions? Whar is the most misleading advice you ever get?
Please write in comments below

Перекласти твіт

5:13 пп · 15 січ. 2022 · Twitter for Android

Переглянути дії з твітом

7 Ретвітів 1 Цитувати твіт 135 Уподобань

Amy Tabb
@amy_tabb

У відповідь @ducha_aiki

Misleading advice: do *whatever* exactly the way someone else does it, whether it be read 10 papers a day, or none.

Your strategy should be allowed to change depending on the needs of the research problem, your own constraints & preferences, etc., and that's ok.

Перекласти твіт

5:27 пп · 15 січ. 2022 · TweetDeck

1 Ретвіт 9 Уподобань

Foreword: trust noone in ML

I mean, really

Foreword: trust noone in ML

I mean, really

- You should develop your own style of reading ML papers, don’t trust me

Foreword: trust noone in ML

I mean, really

- You should develop your own style of reading ML papers, don't trust me
- and never trust authors of the paper

Foreword: trust noone in ML

I mean, really

- You should develop your own style of reading ML papers, don’t trust me
- and never trust authors of the paper
- …even if it was published on CVPR

Foreword: trust noone in ML

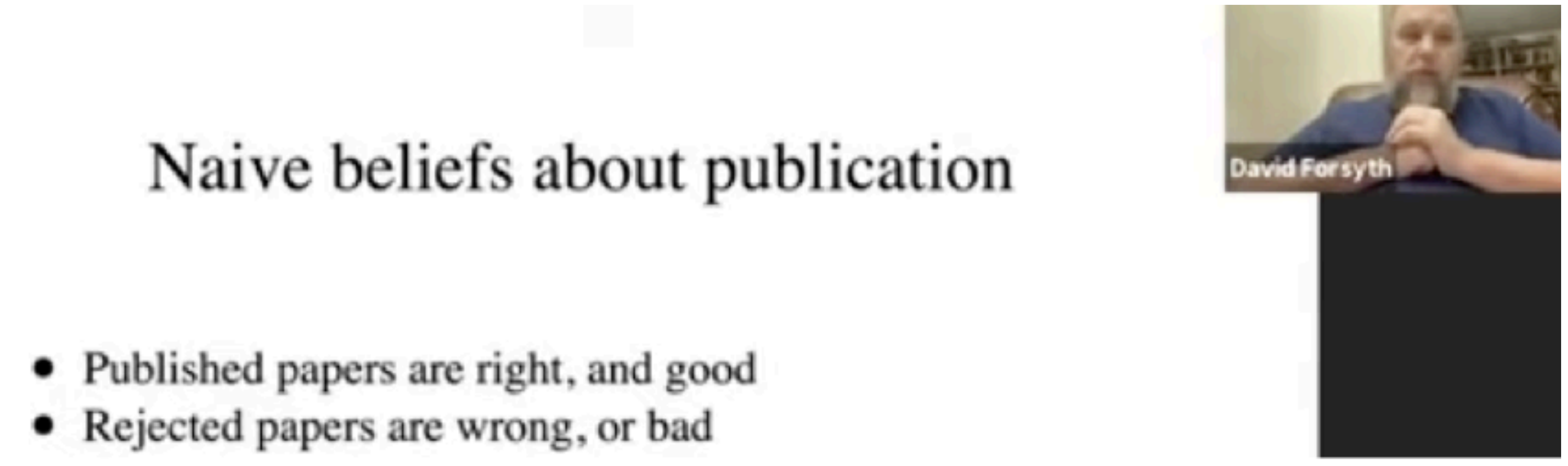
I mean, really

- You should develop your own style of reading ML papers, don’t trust me
- and never trust authors of the paper
- …even if it was published on CVPR
- …especially if it was published on CVPR

Foreword: trust noone in ML

I mean, really

- You should develop your own style of reading ML papers, don’t trust me
- and never trust authors of the paper
- …even if it was published on CVPR
- …especially if it was published on CVPR

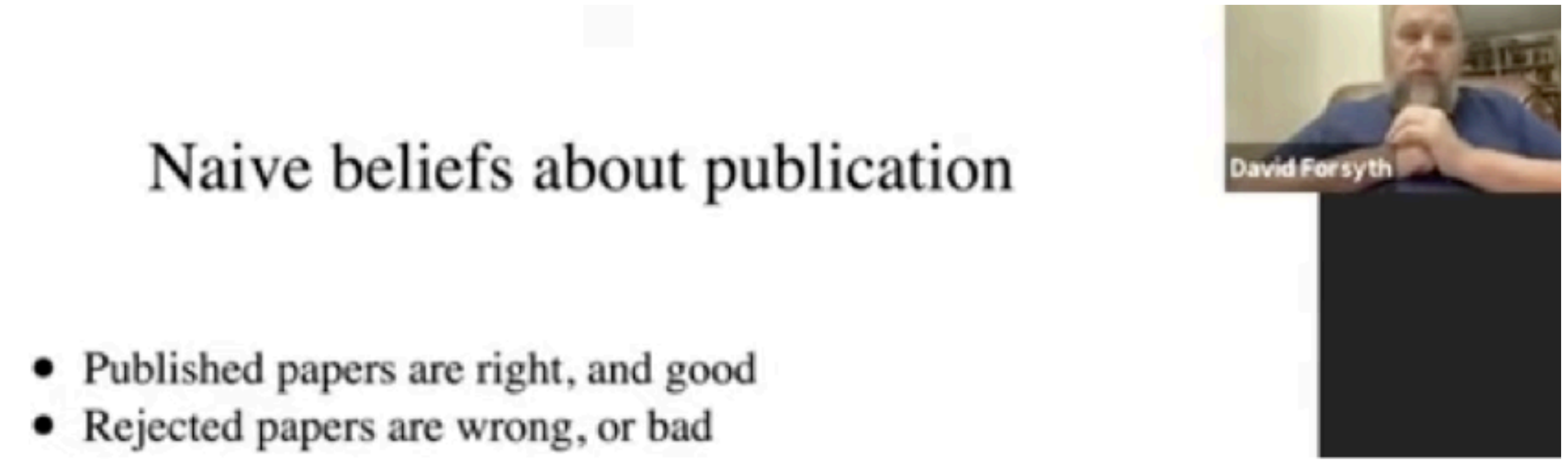


https://www.youtube.com/watch?v=9Y7NCdKdNyE

Foreword: trust noone in ML

I mean, really

- You should develop your own style of reading ML papers, don't trust me
- and never trust authors of the paper
- …even if it was published on CVPR
- …especially if it was published on CVPR
- OK, let's go



https://www.youtube.com/watch?v=9Y7NCdKdNyE

What do I want to achieve?

Why do you need to read papers?

Use-cases

Why do you need to read papers?

Use-cases

(a) You are new to the field and want to get familiar with it

Why do you need to read papers?
Use-cases

(a) You are new to the field and want to get familiar with it

(b) You are new to the field and need to implement something ASAP

Why do you need to read papers?

Use-cases

(a) You are new to the field and want to get familiar with it

(b) You are new to the field and need to implement something ASAP

(c) You are familiar with the field, but not the specific area in particular

Why do you need to read papers?

Use-cases

(a) You are new to the field and want to get familiar with it

(b) You are new to the field and need to implement something ASAP

(c) You are familiar with the field, but not the specific area in particular

(d) You know area well enough and want to keep yourself up-to-date

Why do you need to read papers?

Use-cases

(a) You are new to the field and want to get familiar with it

(b) You are new to the field and need to implement something ASAP

(c) You are familiar with the field, but not the specific area in particular

(d) You know area well enough and want to keep yourself up-to-date

(e) I don’t want, but I have to - I am CVPR reviewer

Why do you need to read papers?
Use-cases

(a) You are new to the field and want to get familiar with it

(b) You are new to the field and need to implement something ASAP

(c) You are familiar with the field, but not the specific area in particular

(d) You know area well enough and want to keep yourself up-to-date

(e) I don’t want, but I have to - I am CVPR reviewer

Each of this cases requires a little bit different strategy

You are new to the field

and need to implement something ASAP (like, yesterday)

- That is really unfortunate situation.
- The only way to succeed is to be lucky and find ready implementation
- Check the following first:
  - Kaggle & DrivenData competitions
  - Papers With Code
  - fast.ai forums, pyimagesearch, etc.
  - Only then browse papers and github.

Kaggle

https://www.kaggle.com

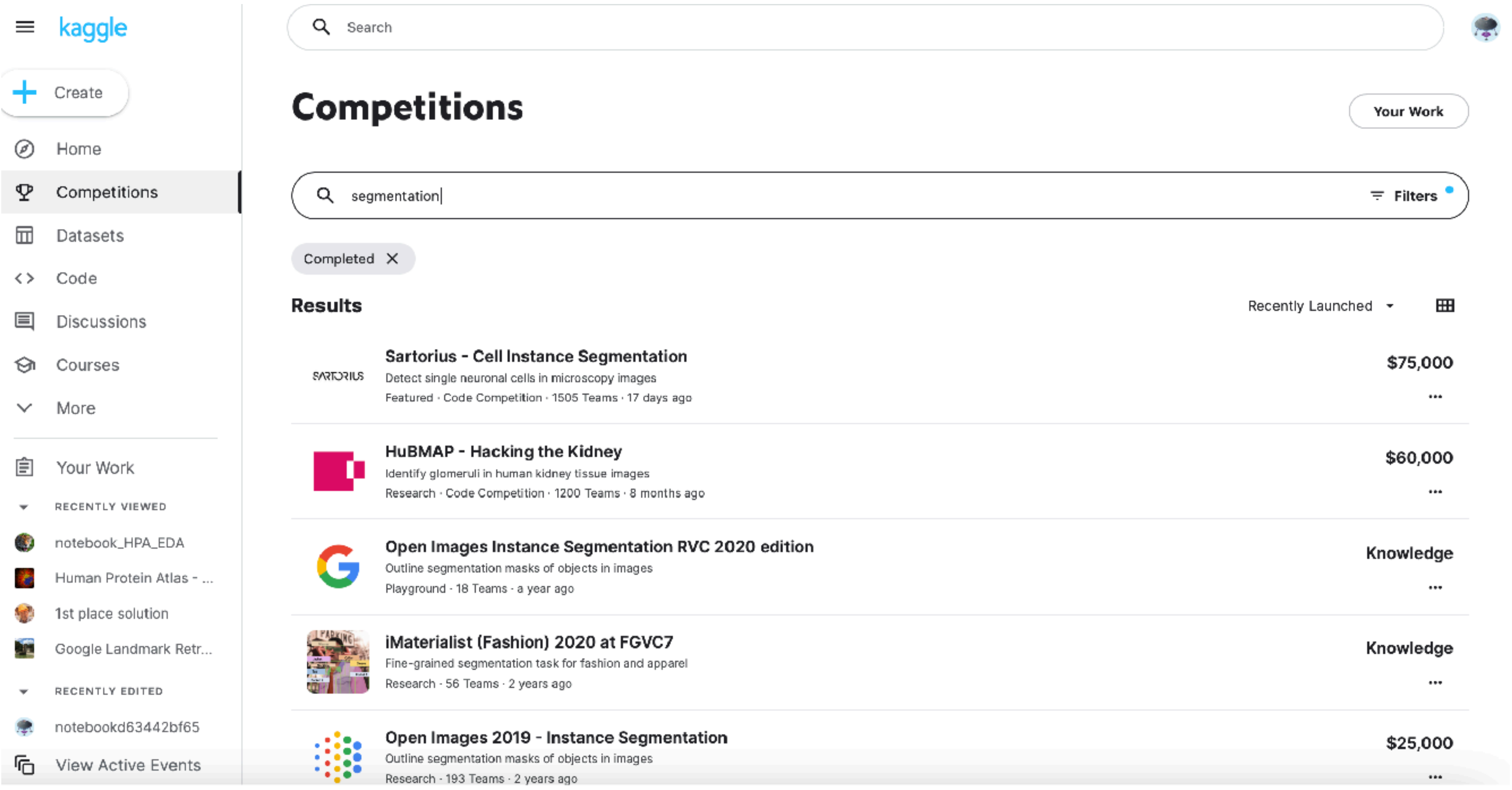
Kaggle

Competitions → keyword search → EDA kernels

https://www.kaggle.com

Kaggle

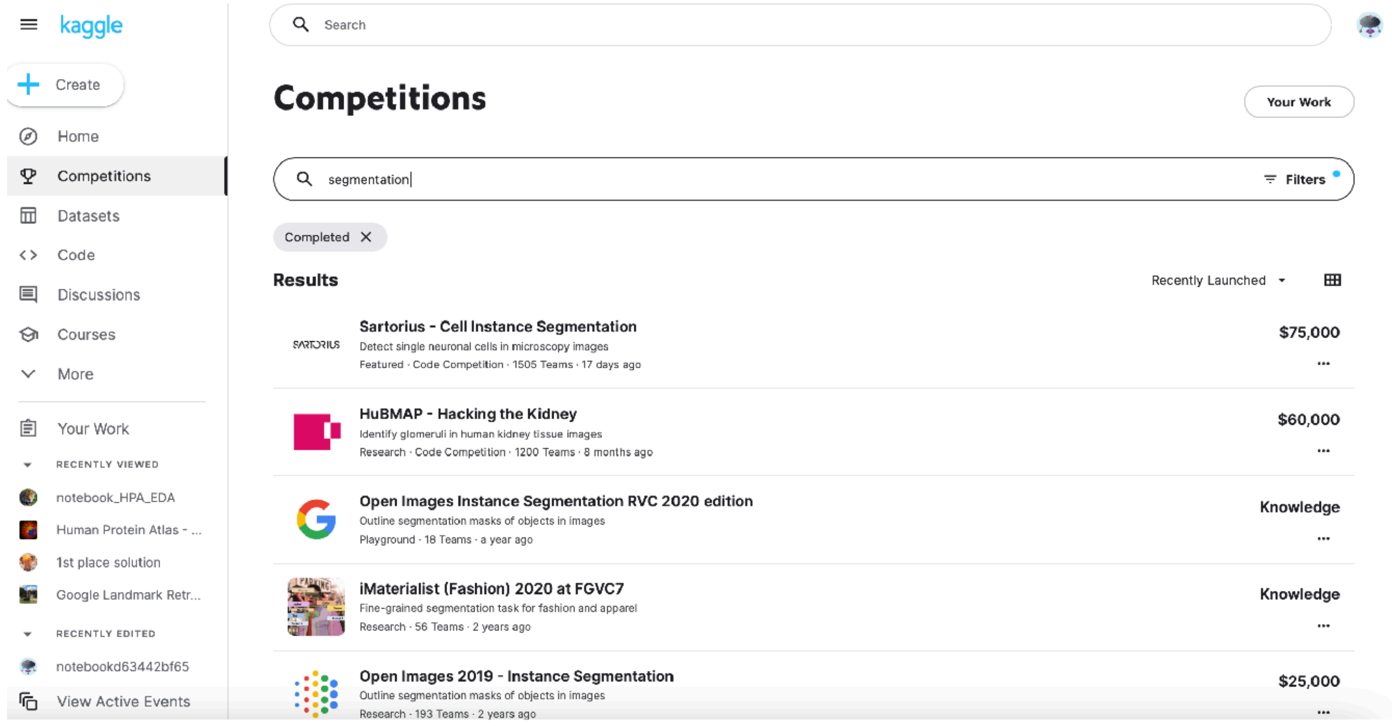
Competitions → keyword search → EDA kernels



https://www.kaggle.com

Kaggle

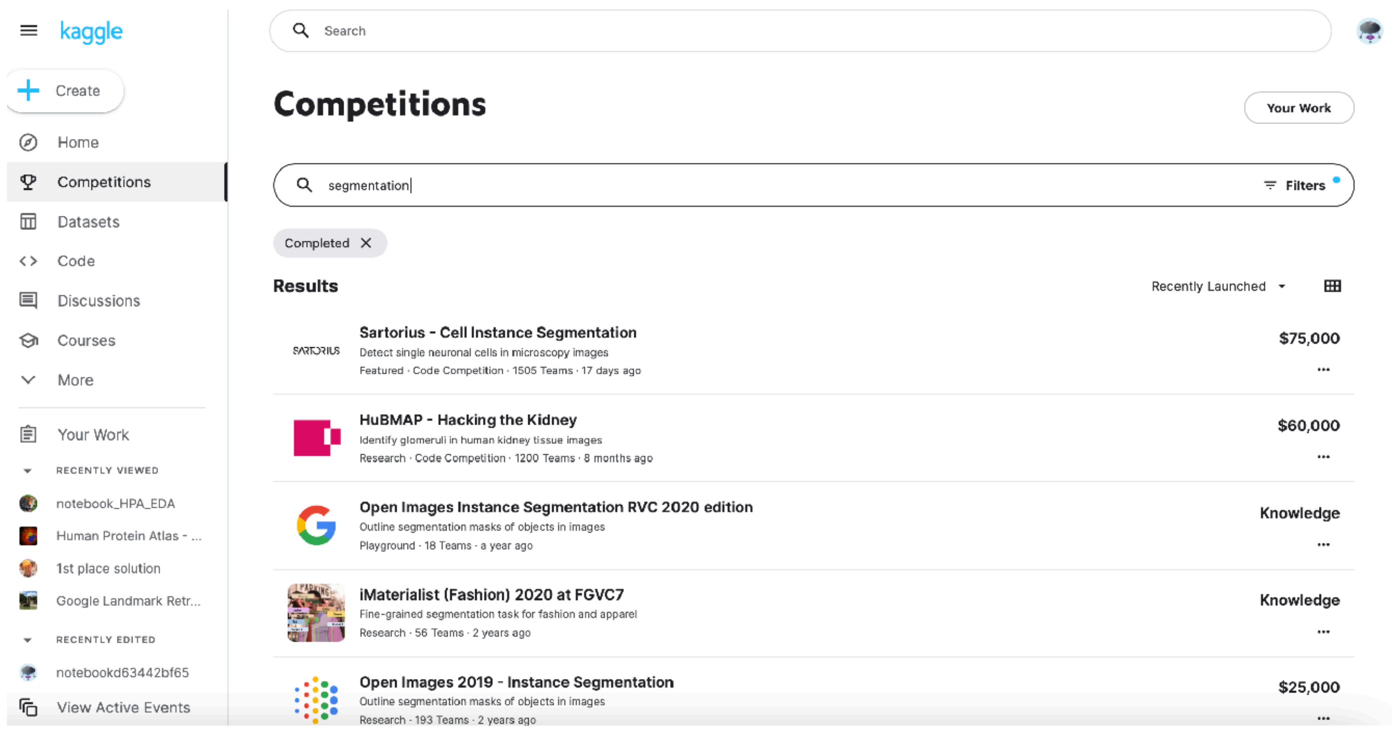
Competitions → keyword search → EDA kernels

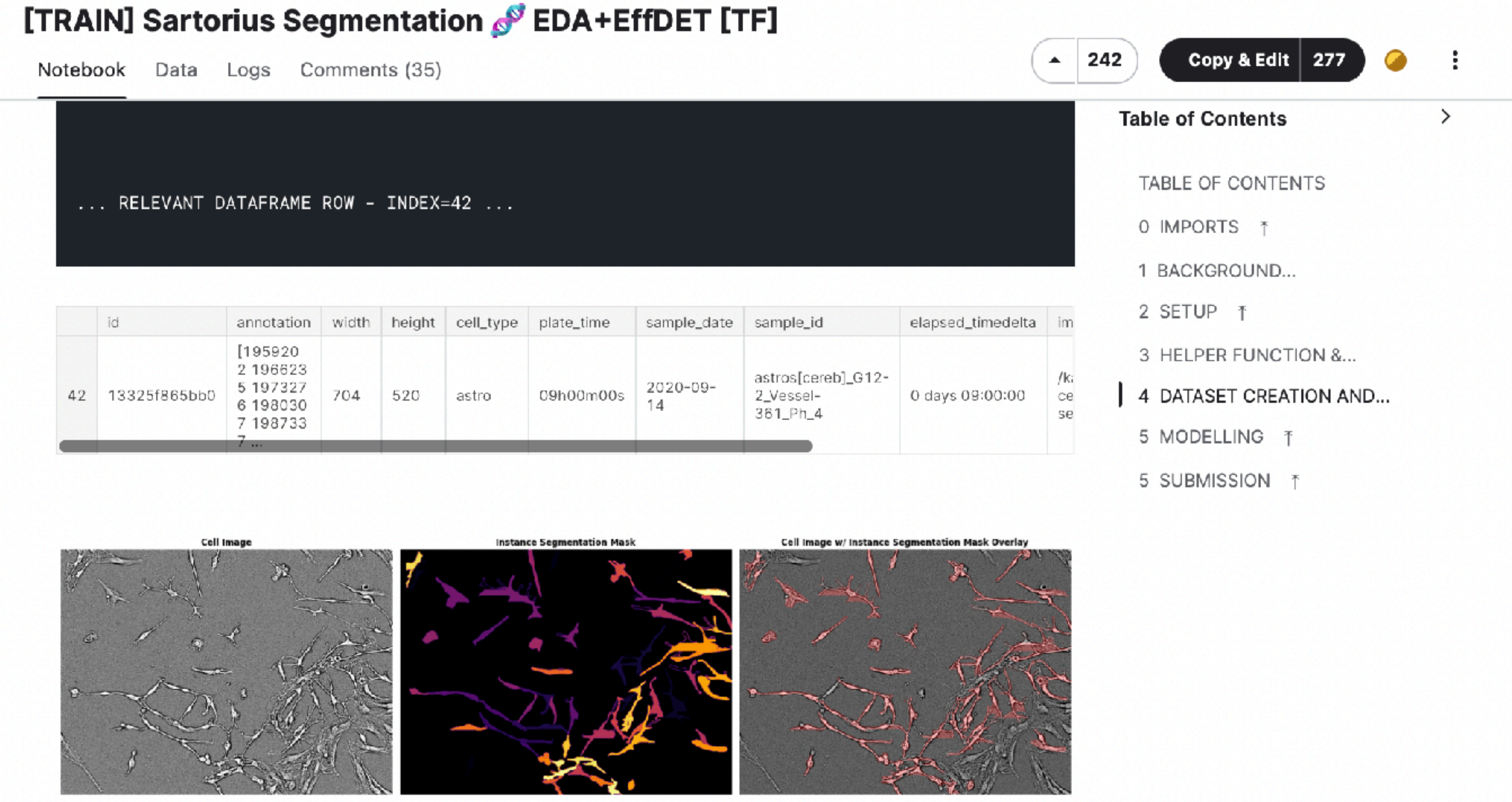


https://www.kaggle.com

Kaggle

Competitions → keyword search → EDA kernels



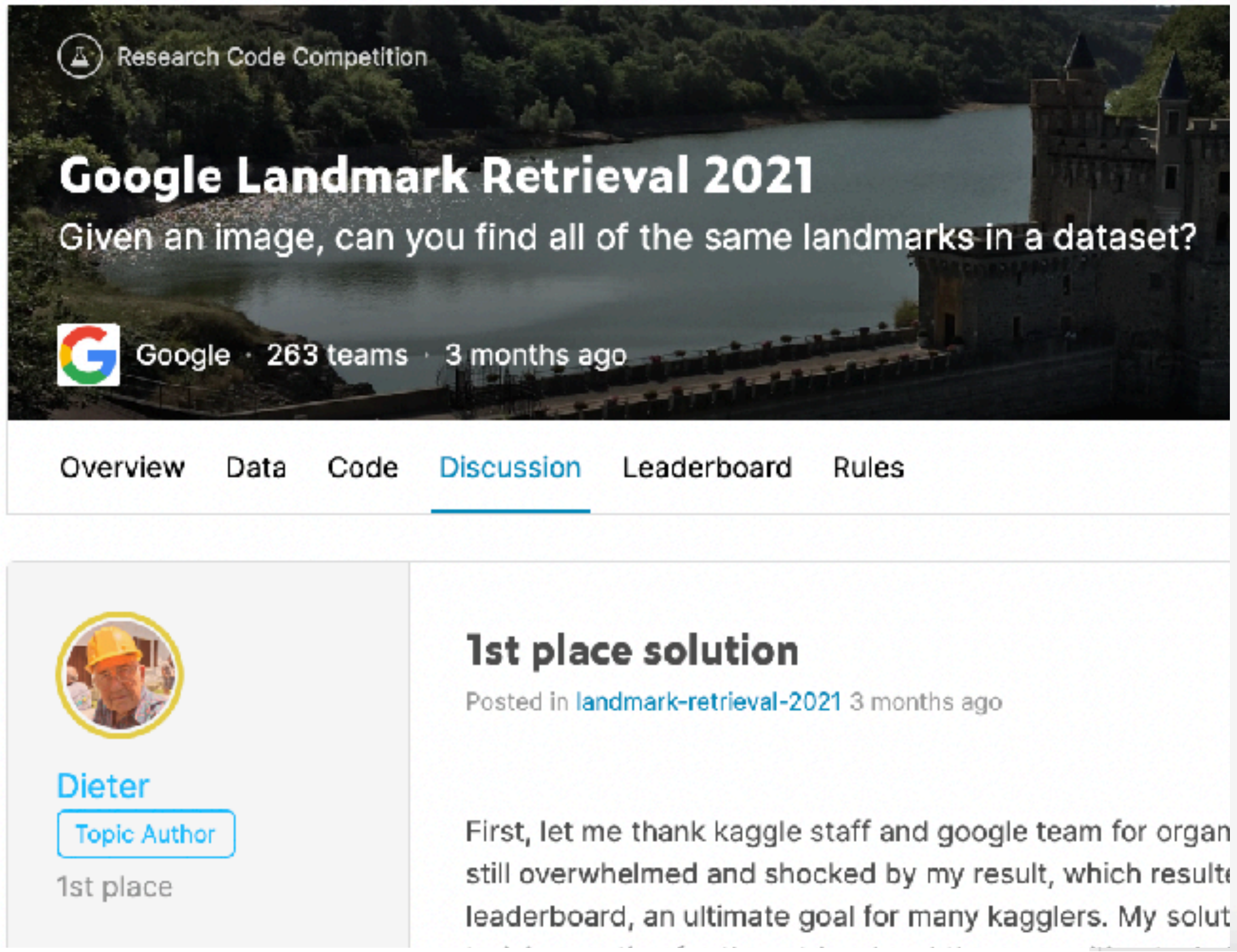


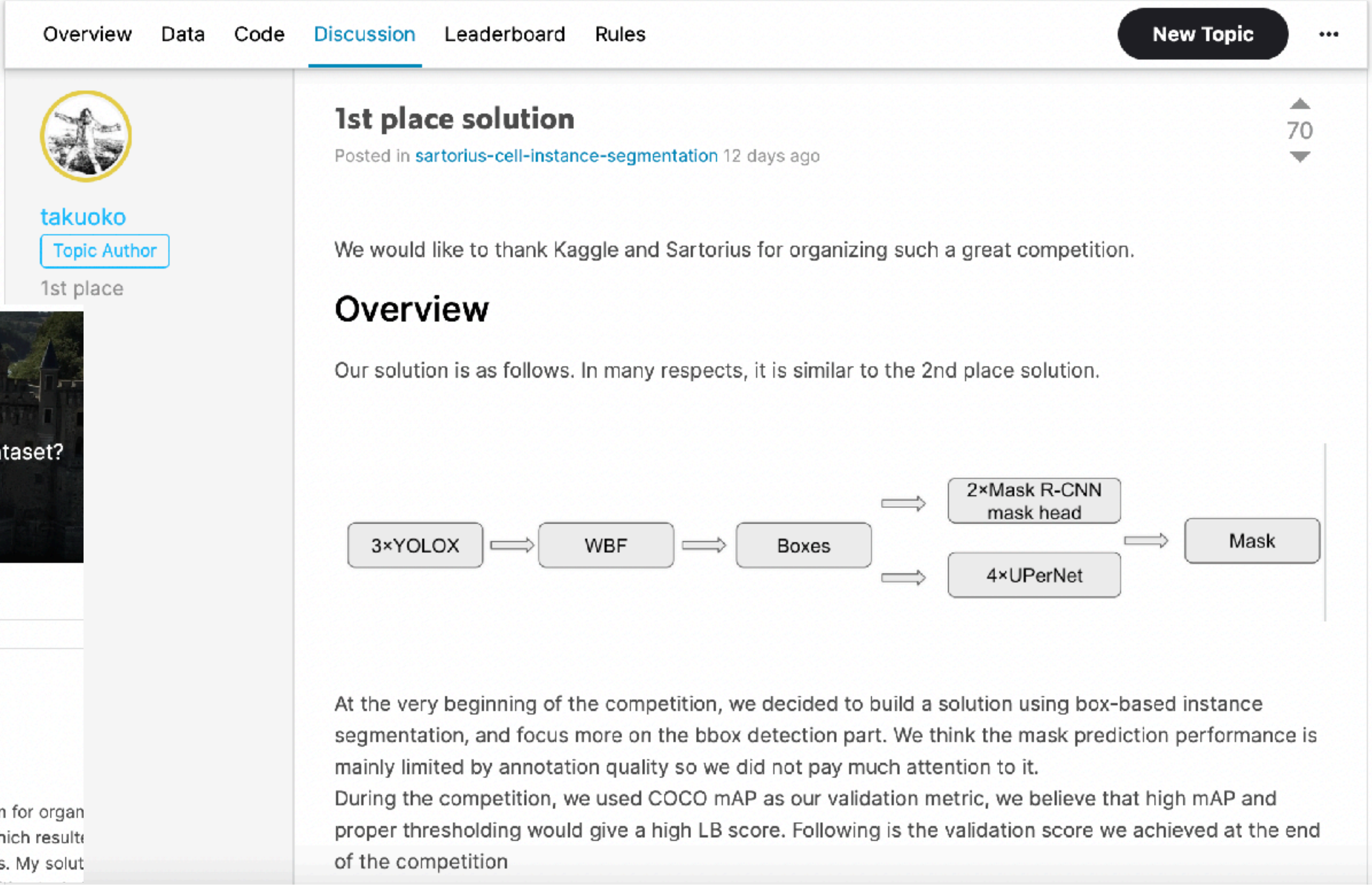
https://www.kaggle.com

Kaggle

If task & competition is relevant to you, check solutions

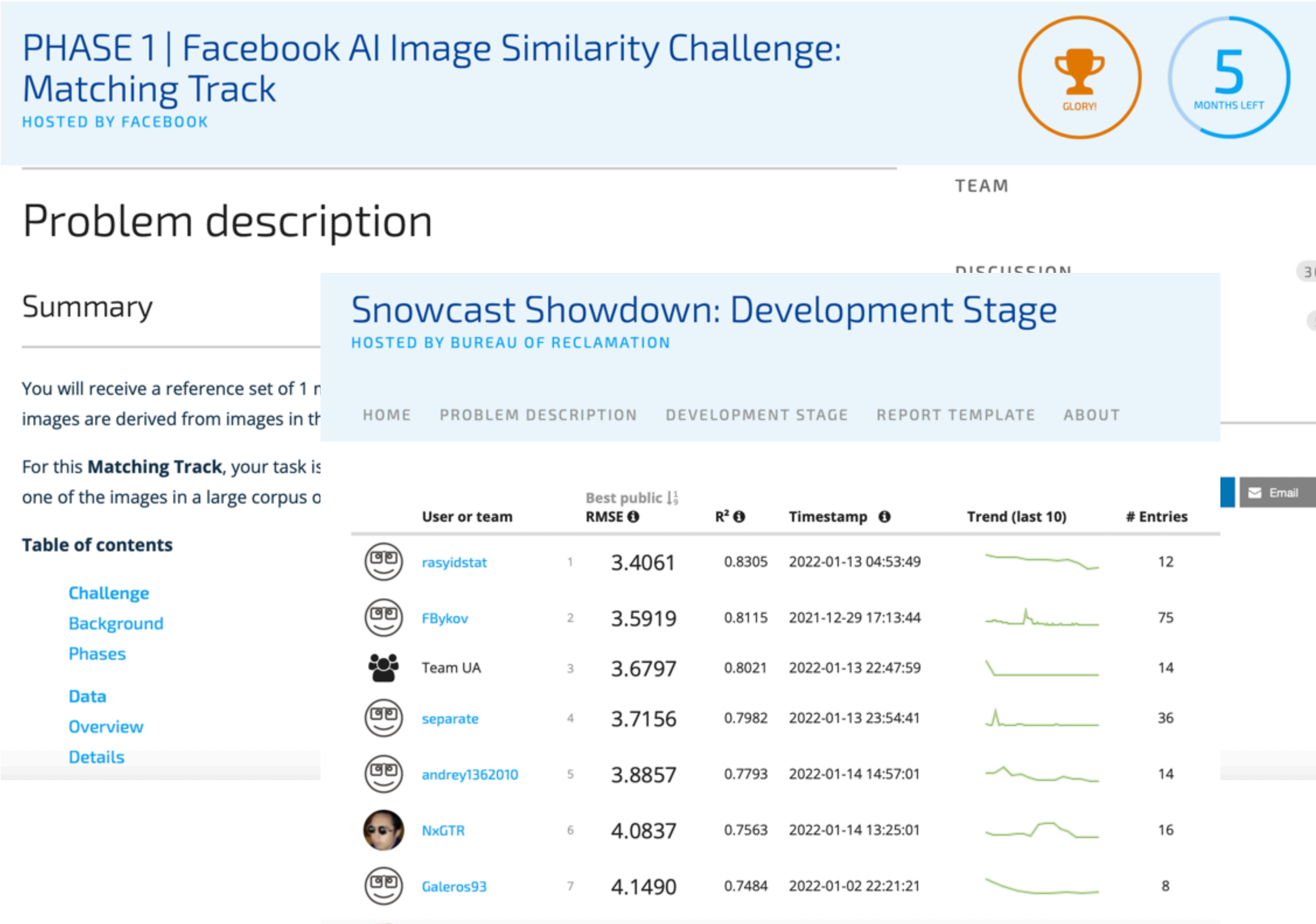
- Winning solutions
- Simple solutions





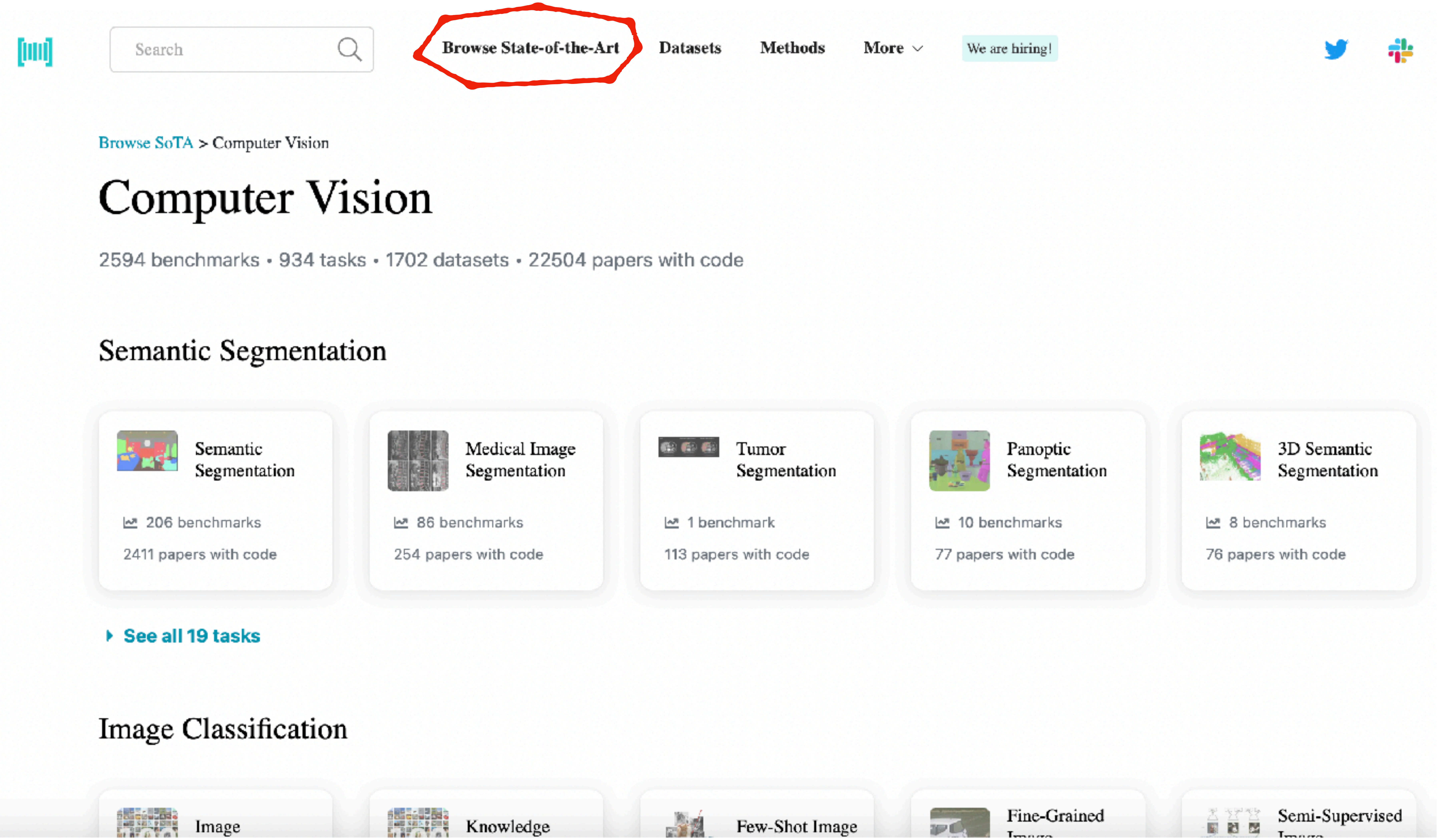
DrivenData

- Much smaller community than Kaggle, with much less forum life
- But contains some unique challenges



https://www.kaggle.com

Papers With Code



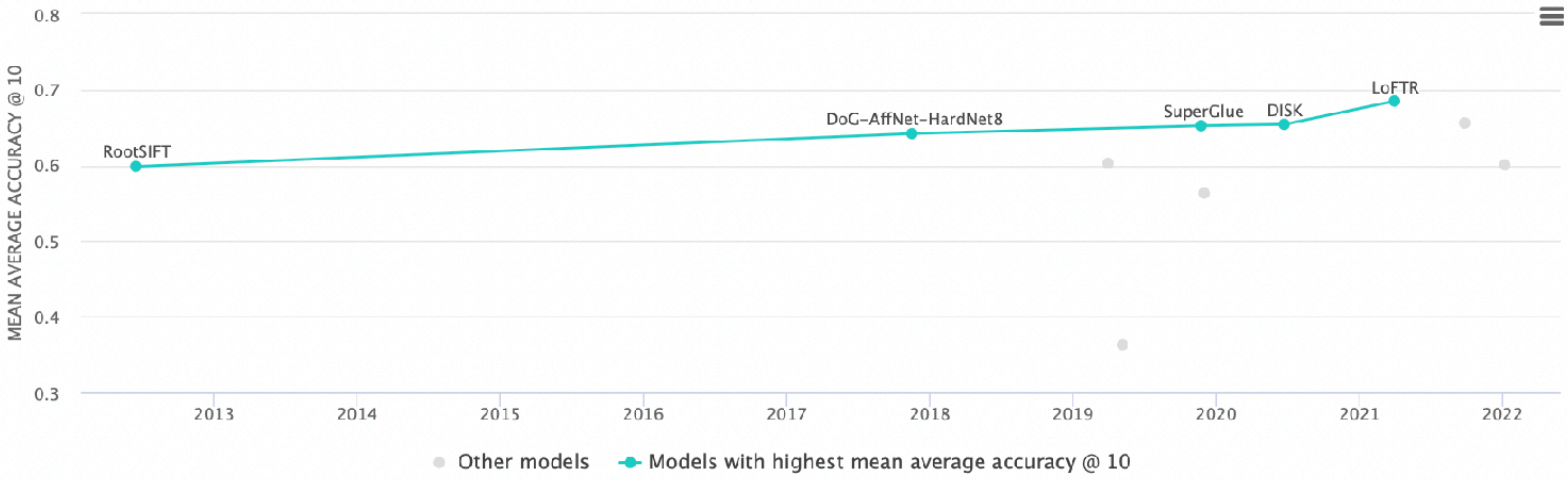
Papers With Code

Caveat: sparsely populated

Image Matching on IMC PhotoTourism

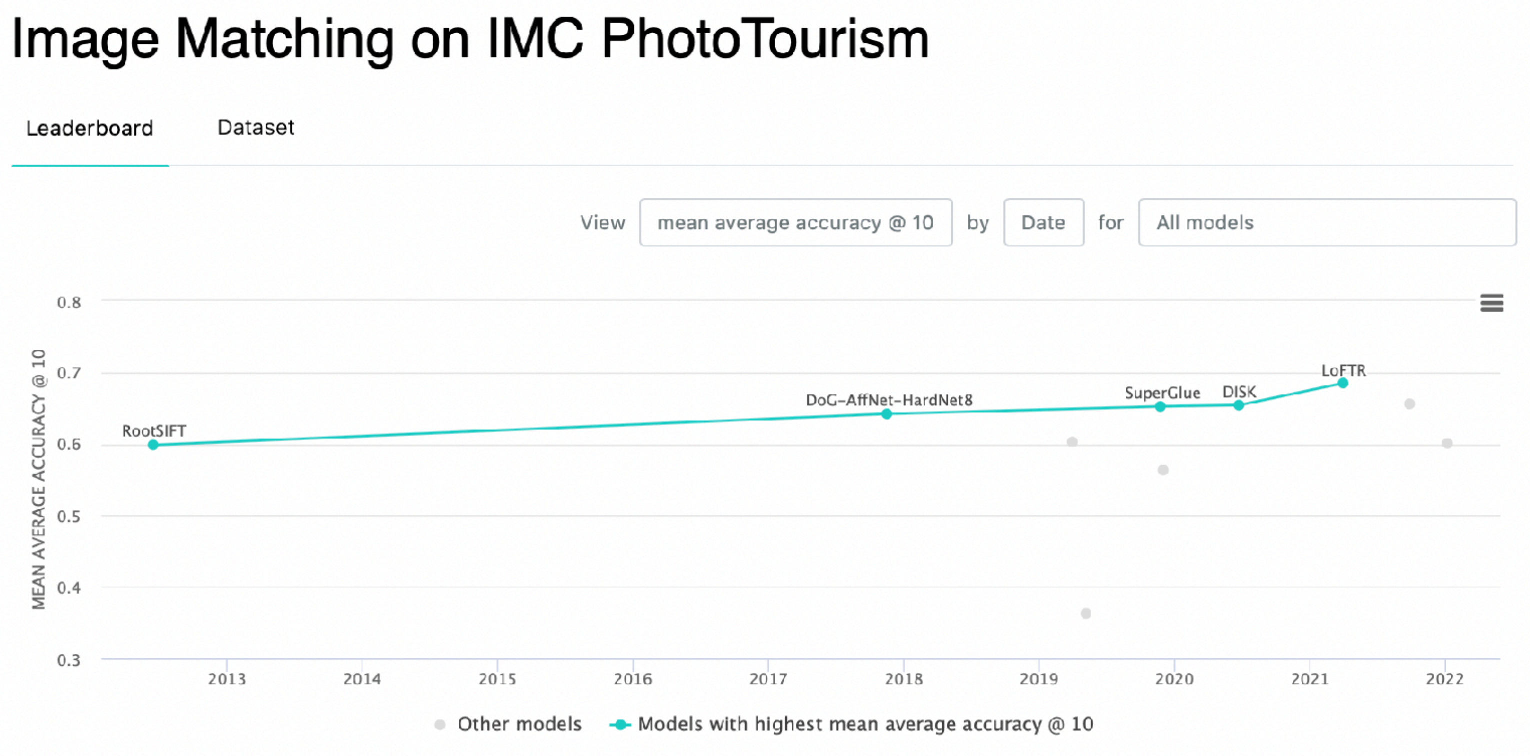
Leaderboard Dataset

View mean average accuracy @ 10 by Date for All models



Papers With Code

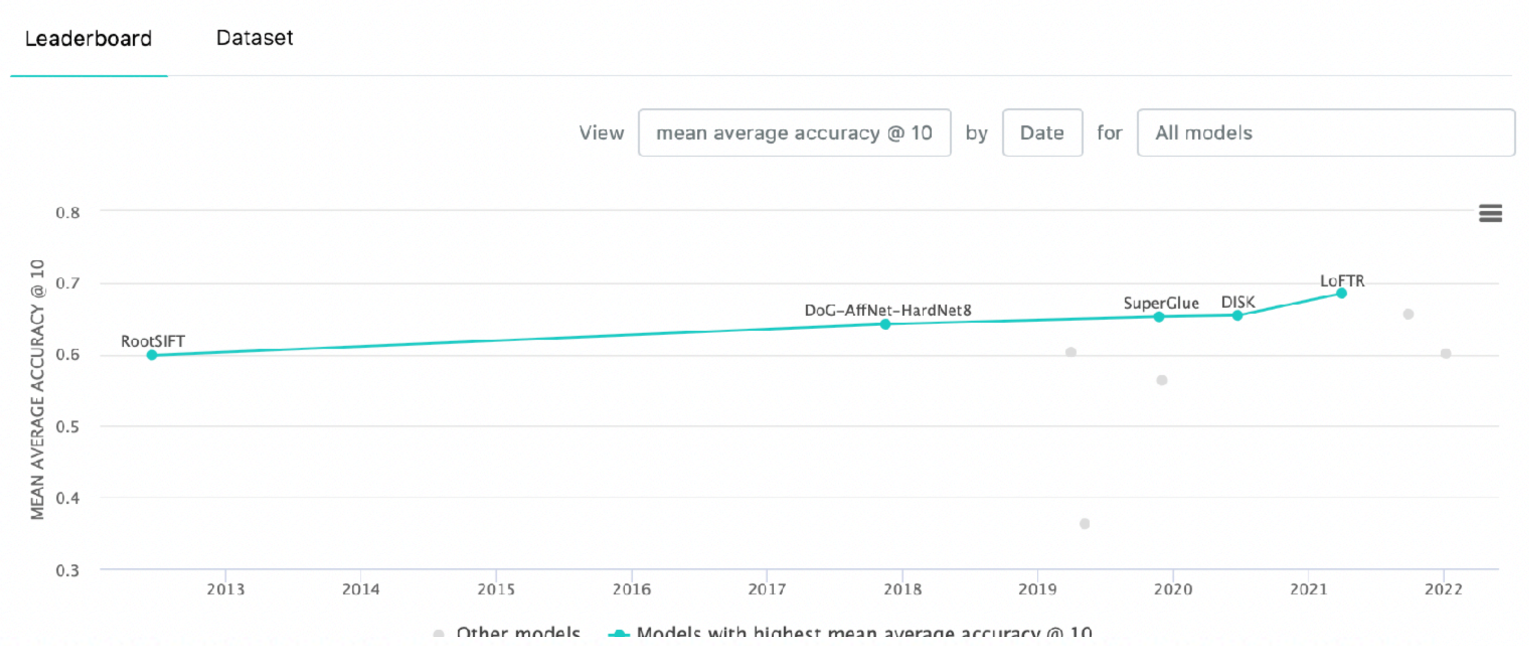
Caveat: sparsely populated



Papers With Code

Caveat: sparsely populated

Image Matching on IMC PhotoTourism



Rank	Model	mean average accuracy @ 10	Extra Training Data	Paper	Code	Result	Year	Tags
1	**LoFTR**	0.68503	✓	LoFTR: Detector-Free Local Feature Matching with Transformers			2021	
2	**HarrisZ improved, Blob DTM**	0.65606	✗	HarrisZ^{+}: Harris Corner Selection for Next-Gen Image Matching Pipelines			2021	
3	**DISK**	0.65435	✓	DISK: Learning local features with policy gradient			2020	
4	**SuperGlue**	0.65248	✓	SuperGlue: Learning Feature Matching with Graph Neural Networks			2019	
5	**DoG-AffNet-HardNet8**	0.64212	✓	Repeatability Is Not Enough: Learning Affine Regions via Discriminability			2017	
6	**Key.Net-SOSNet**	0.60285	✗	Key.Net: Keypoint Detection by Handcrafted and Learned CNN Filters			2019	
7	**PoSFeat**	0.60072	✗	Decoupling Makes Weakly Supervised Local Feature Better			2022	
8	**RootSIFT**	0.59859	✗	Three things everyone should know to improve object retrieval			2012	
9	**R2D2**	0.56345	✓	R2D2: Reliable and Repeatable Detector and Descriptor			2019	

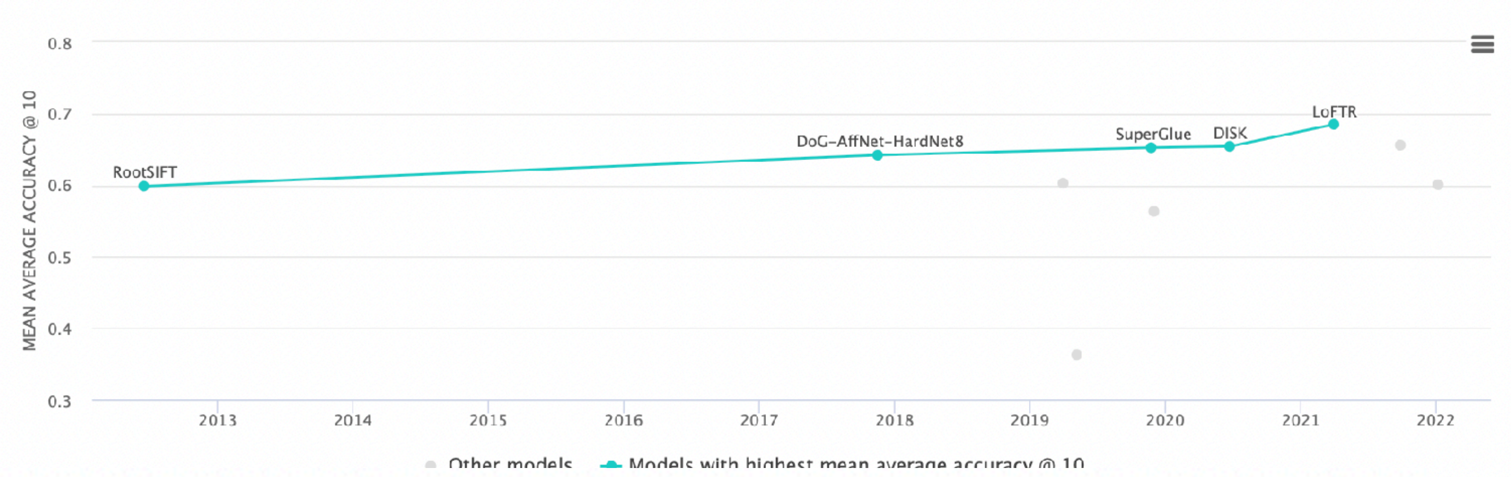
Papers With Code

Caveat: sparsely populated

Image Matching on IMC PhotoTourism

Leaderboard Dataset

View mean average accuracy @ 10 by Date for All models



I have added this whole leaderboard

when preparing for this talk.

Before there was no leaderboard, and nobody checked it

Rank	Model	mean average accuracy @ 10	Extra Training Data	Paper	Code	Result	Year	Tags
1	LoFTR	0.68503	✓	LoFTR: Detector-Free Local Feature Matching with Transformers			2021	
2	HarrisZ improved, Blob DTM	0.65606	×	HarrisZ^{+}: Harris Corner Selection for Next-Gen Image Matching Pipelines			2021	
3	DISK	0.65435	✓	DISK: Learning local features with policy gradient			2020	
4	SuperGlue	0.65248	✓	SuperGlue: Learning Feature Matching with Graph Neural Networks			2019	
5	DoG-AffNet-HardNet8	0.64212	✓	Repeatability Is Not Enough: Learning Affine Regions via Discriminability			2017	
6	Key.Net-SOSNet	0.60285	×	Key.Net: Keypoint Detection by Handcrafted and Learned CNN Filters			2019	
7	PoSFeat	0.60072	×	Decoupling Makes Weakly Supervised Local Feature Better			2022	
8	RootSIFT	0.59859	×	Three things everyone should know to improve object retrieval			2012	
9	R2D2	0.56345	✓	R2D2: Reliable and Repeatable Detector and Descriptor			2019	

Papers With Code

Caveat: sparsely populated

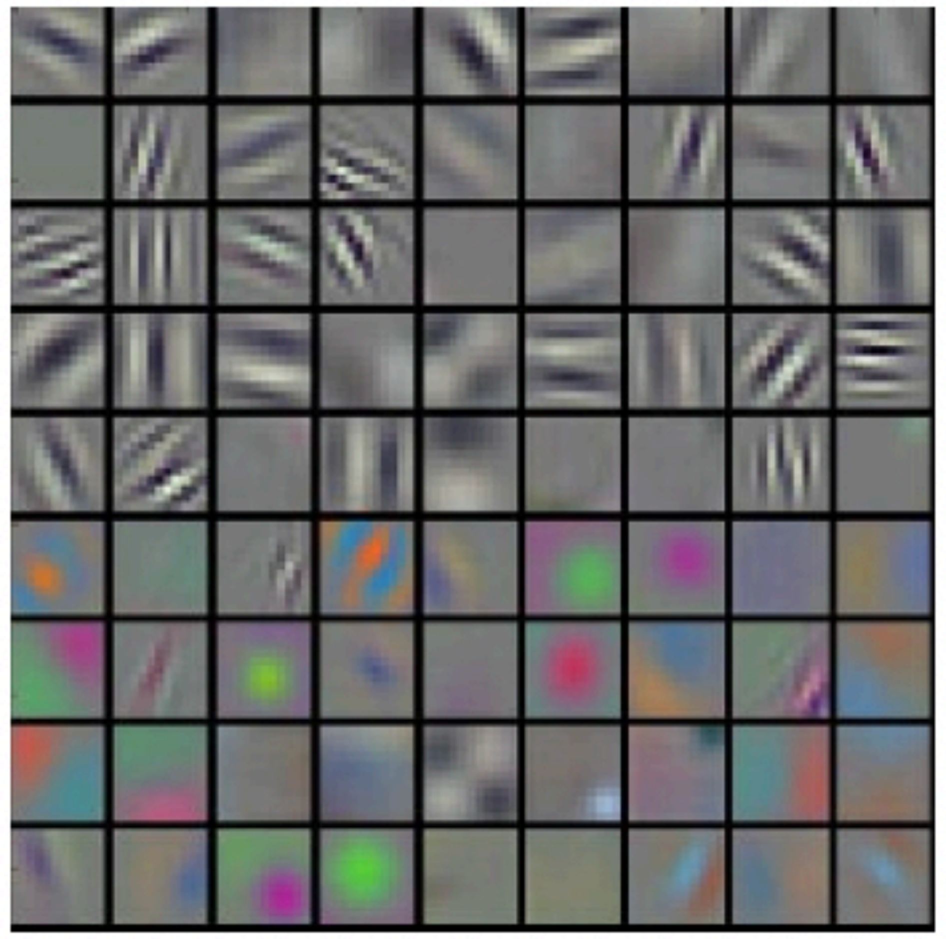
Methodology

Representation Learning

Edit

1643 papers with code • 2 benchmarks • 2 datasets

Representation learning is concerned with training machine learning algorithms to learn useful representations, e.g. those that are interpretable, have latent features, or can be used for transfer learning.



Deep neural networks can be considered representation learning models that typically encode information which is projected into a different subspace. These representations are then usually passed on to a linear classifier to, for instance, train a classifier.

Representation learning can be divided into:

- **Supervised representation learning**: learning representations on task A using annotated data and used to solve task B

Content

Papers With Code

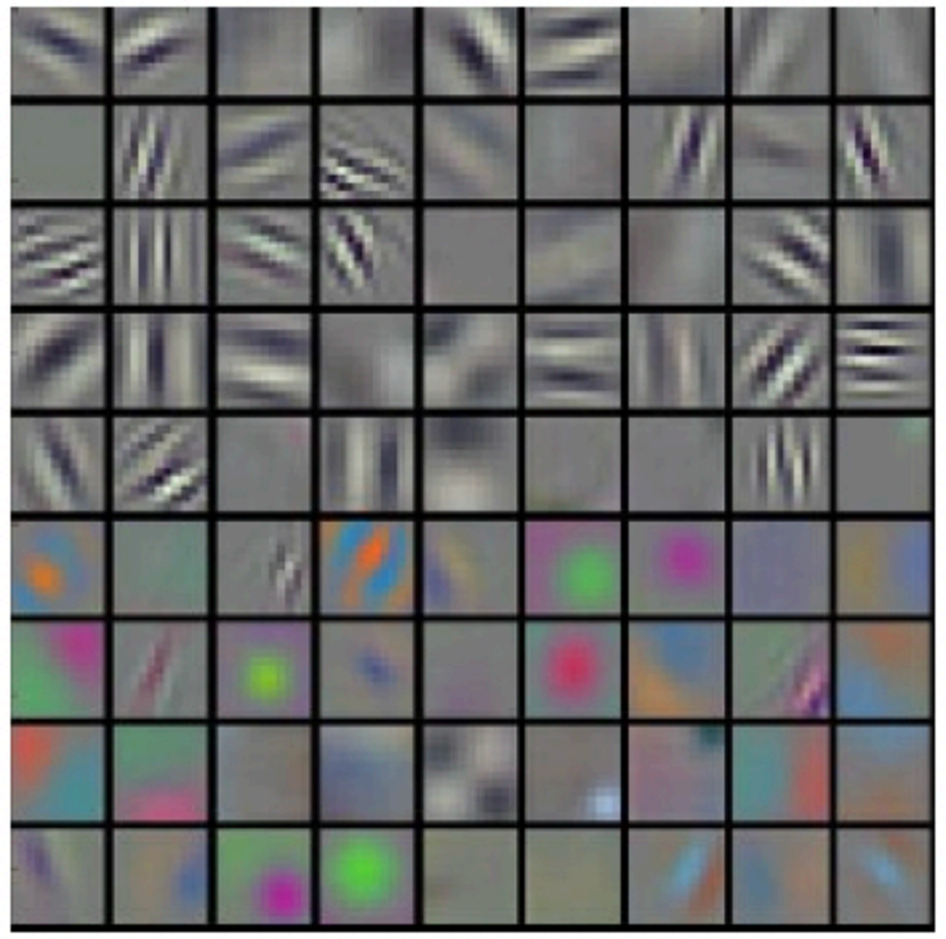
Caveat: sparsely populated

Methodology

Representation Learning

Edit

1643 papers with code • 2 benchmarks • 2 datasets



Representation learning is concerned with training machine learning algorithms to learn useful representations, e.g. those that are interpretable, have latent features, or can be used for transfer learning.

Deep neural networks can be considered representation learning models that typically encode information which is projected into a different subspace. These representations are then usually passed on to a linear classifier to, for instance, train a classifier.

Representation learning can be divided into:

- **Supervised representation learning**: learning representations on task A using annotated data and used to solve task B

Content

Papers With Code

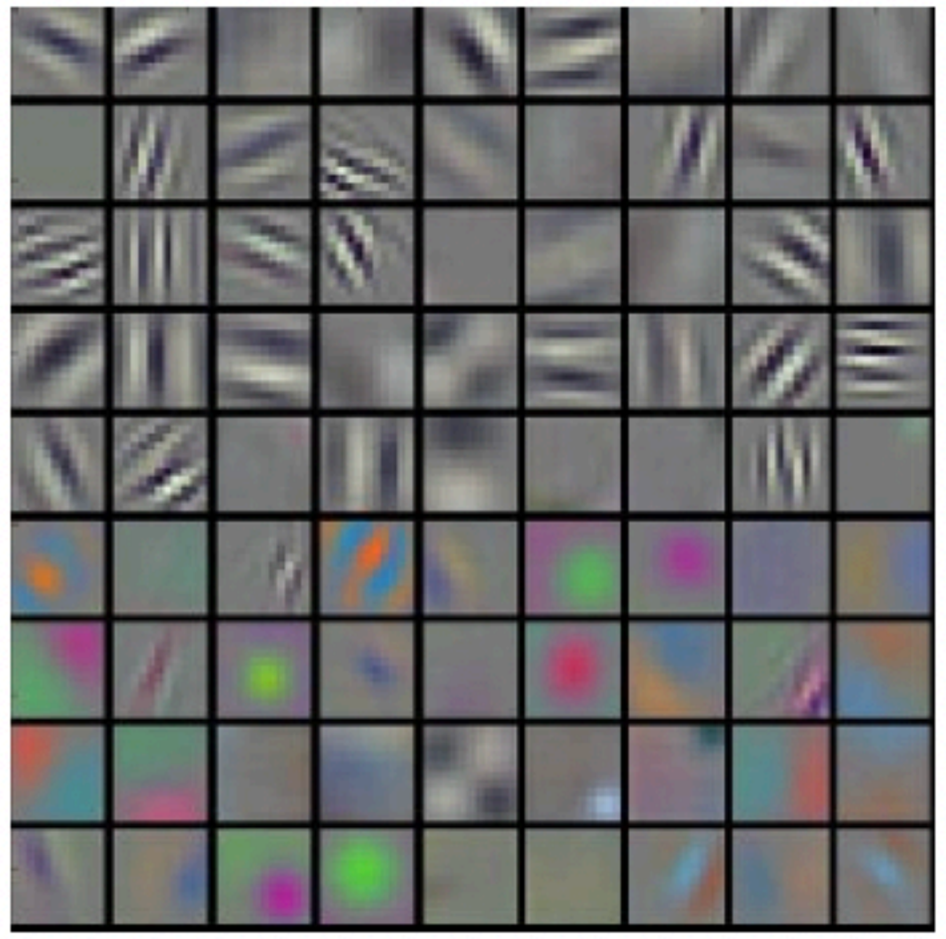
Caveat: sparsely populated

Methodology

Representation Learning

1643 papers with code • 2 benchmarks • 2 datasets

Representation learning is concerned with training machine learning algorithms to learn useful representations, e.g. those that are interpretable, have latent features, or can be used for transfer learning.



Content

Deep neural networks can be considered representation learning models that typically encode information which is projected into a different subspace. These representations are then usually passed on to a linear classifier to, for instance, train a classifier.

Representation learning can be divided into:

- **Supervised representation learning**: learning representations on task A using annotated data and used to solve task B

Papers With Code

Caveat: sparsely populated

Methodology

Representation Learning

Edit

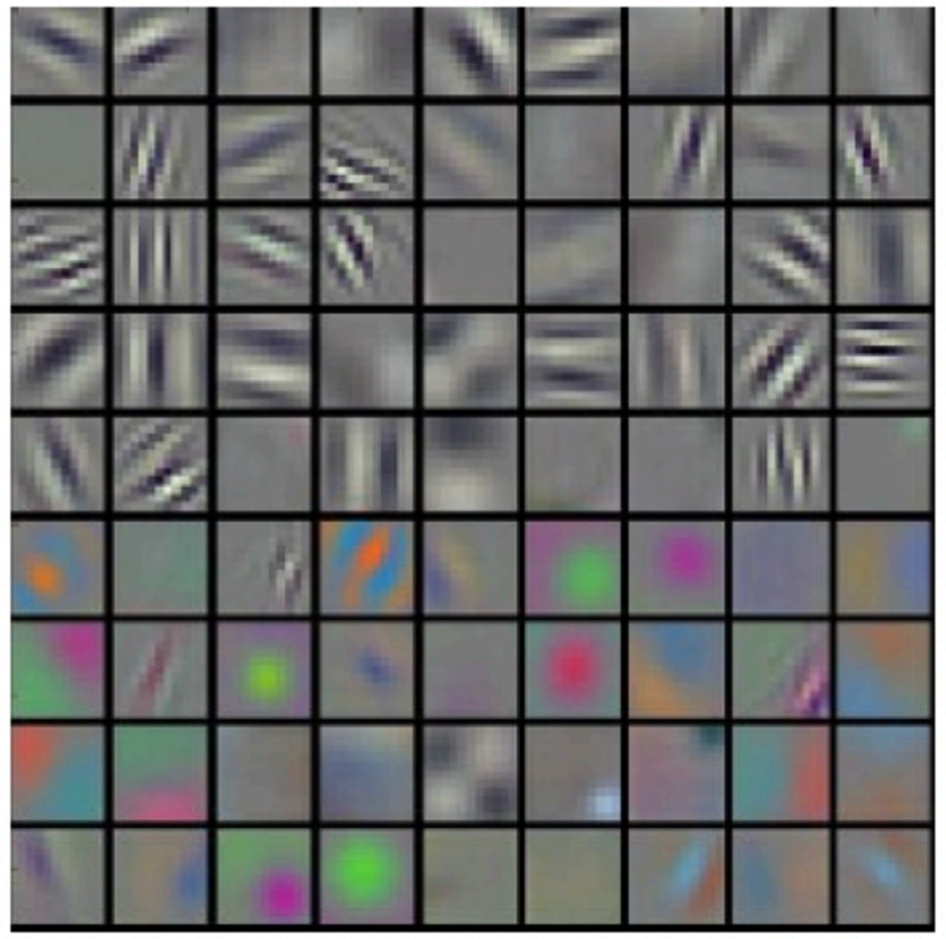
1643 papers with code • 2 benchmarks • 2 datasets

Representation learning is concerned with training machine learning algorithms to learn useful representations, e.g. those that are interpretable, have latent features, or can be used for transfer learning.

Deep neural networks can be considered representation learning models that typically encode information which is projected into a different subspace. These representations are then usually passed on to a linear classifier to, for instance, train a classifier.

Representation learning can be divided into:

- **Supervised representation learning**: learning representations on task A using annotated data and used to solve task B



Content

Papers With Code
Caveat: sparsely populated

Methodology

Edit

Representation Learning

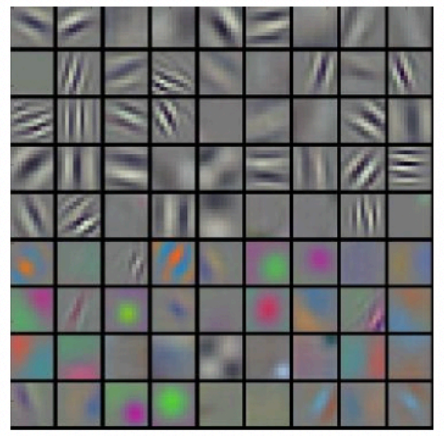
1643 papers with code • 2 benchmarks • 2 datasets

Representation learning is concerned with training machine learning algorithms to learn useful representations, e.g. those that are interpretable, have latent features, or can be used for transfer learning.

Deep neural networks can be considered representation learning models that typically encode information which is projected into a different subspace. These representations are then usually passed on to a linear classifier to, for instance, train a classifier.

Representation learning can be divided into:

- **Supervised representation learning**: learning representations on task A using annotated data and used to solve task B



Content

Papers With Code

Caveat: sparsely populated

Methodology

Representation Learning

1643 papers with code · 2 benchmarks · 2 datasets

Representation learning is concerned with training machine learni representations, e.g. those that are interpretable, have latent feature learning.

Deep neural networks can be considered representation learning information which is projected into a different subspace. These re passed on to a linear classifier to, for instance, train a classifier.

Representation learning can be divided into:

- **Supervised representation learning**: learning representations on and used to solve task B

Benchmarks

Add a Result

These leaderboards are used to track progress in Representation Learning

Trend	Dataset	Best Model	Paper	Code	Compare
	Circle Data	Morphological Network			See all
	Sports10	Max Margin Contrastive			See all

Papers With Code

Caveat: sparsely populated

Methodology

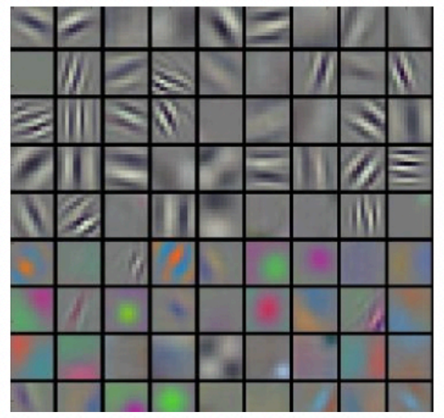
Representation Learning

Edit

1643 papers with code · 2 benchmarks · 2 datasets

Representation learning is concerned with training machine learning algorithms to learn useful representations, e.g. those that are interpretable, have latent features, or can be used for transfer learning.

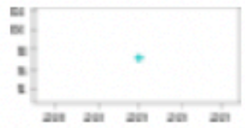
Deep neural networks can be considered representation learning models that typically encode information which is projected into a different subspace. These representations are then usually passed on to a linear classifier to, for instance, train a classifier.



Benchmarks

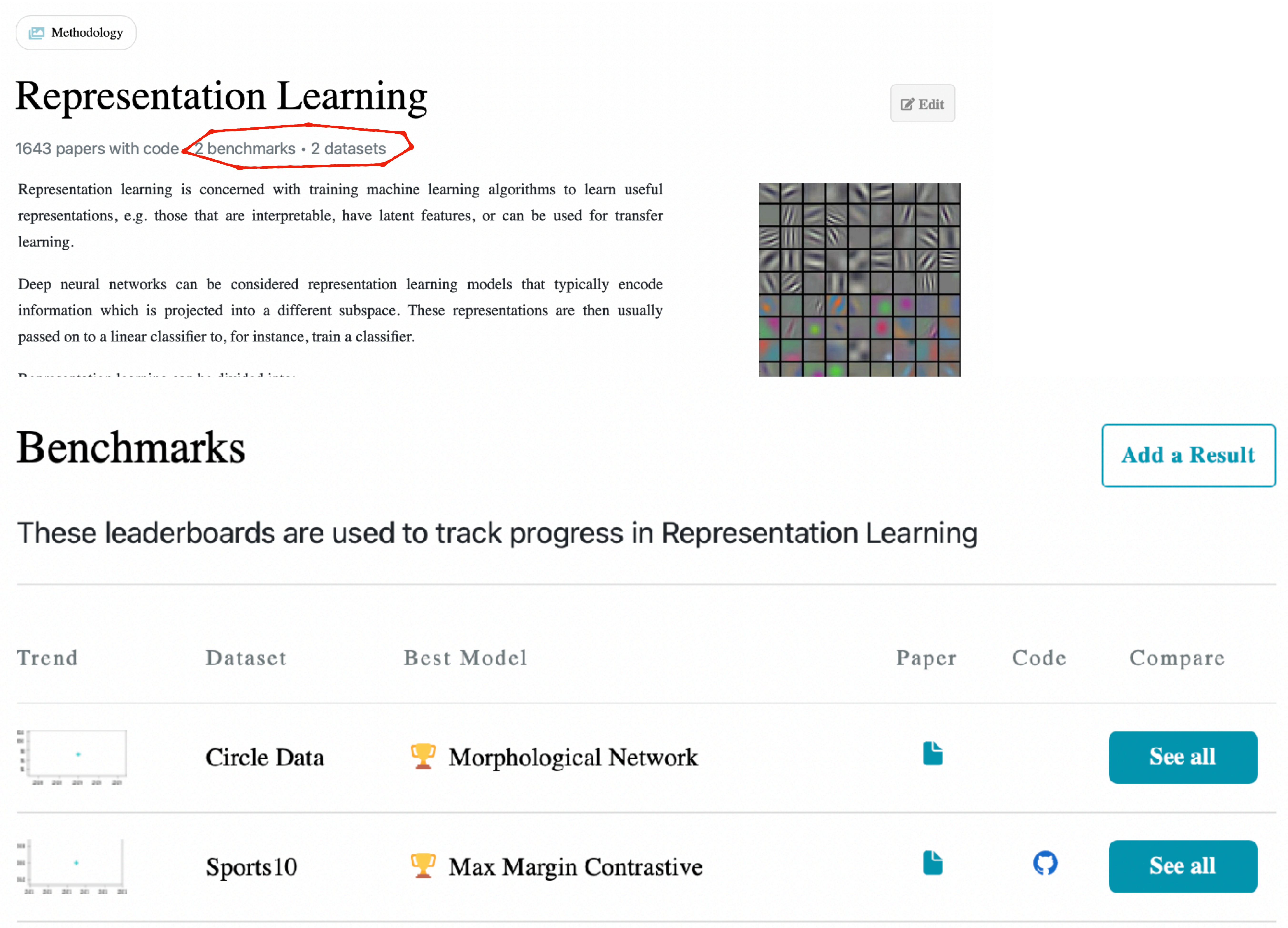
Add a Result

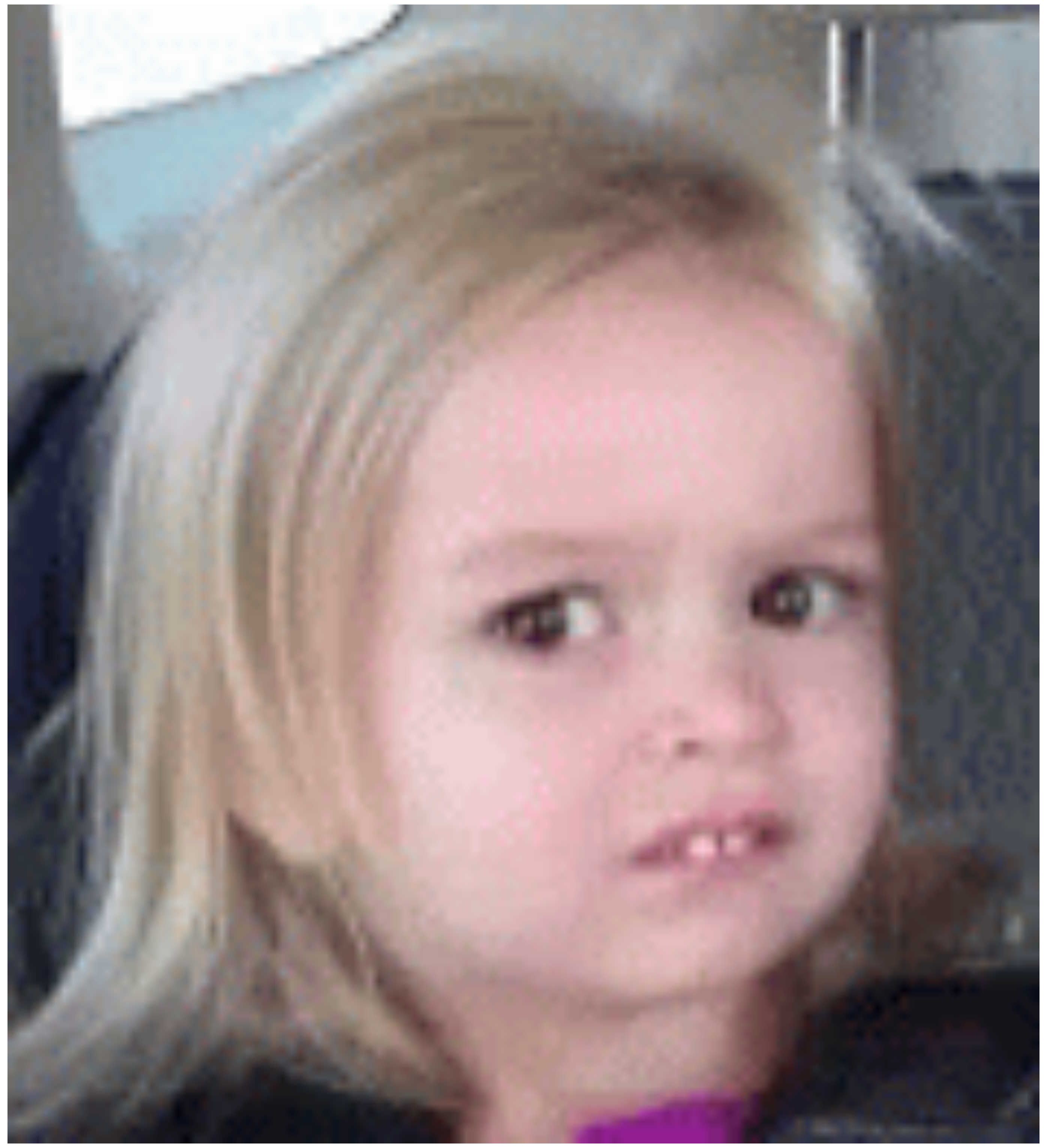
These leaderboards are used to track progress in Representation Learning

Trend	Dataset	Best Model	Paper	Code	Compare
	Circle Data	Morphological Network			See all
	Sports10	Max Margin Contrastive			See all

Papers With Code

Caveat: sparsely populated





If both Kaggle & PapersWithCode fail

Then go to github + arXiv search

If both Kaggle & PapersWithCode fail

Then go to github + arXiv search

- You might want to apply heuristics to decide which papers you should read first

If both Kaggle & PapersWithCode fail

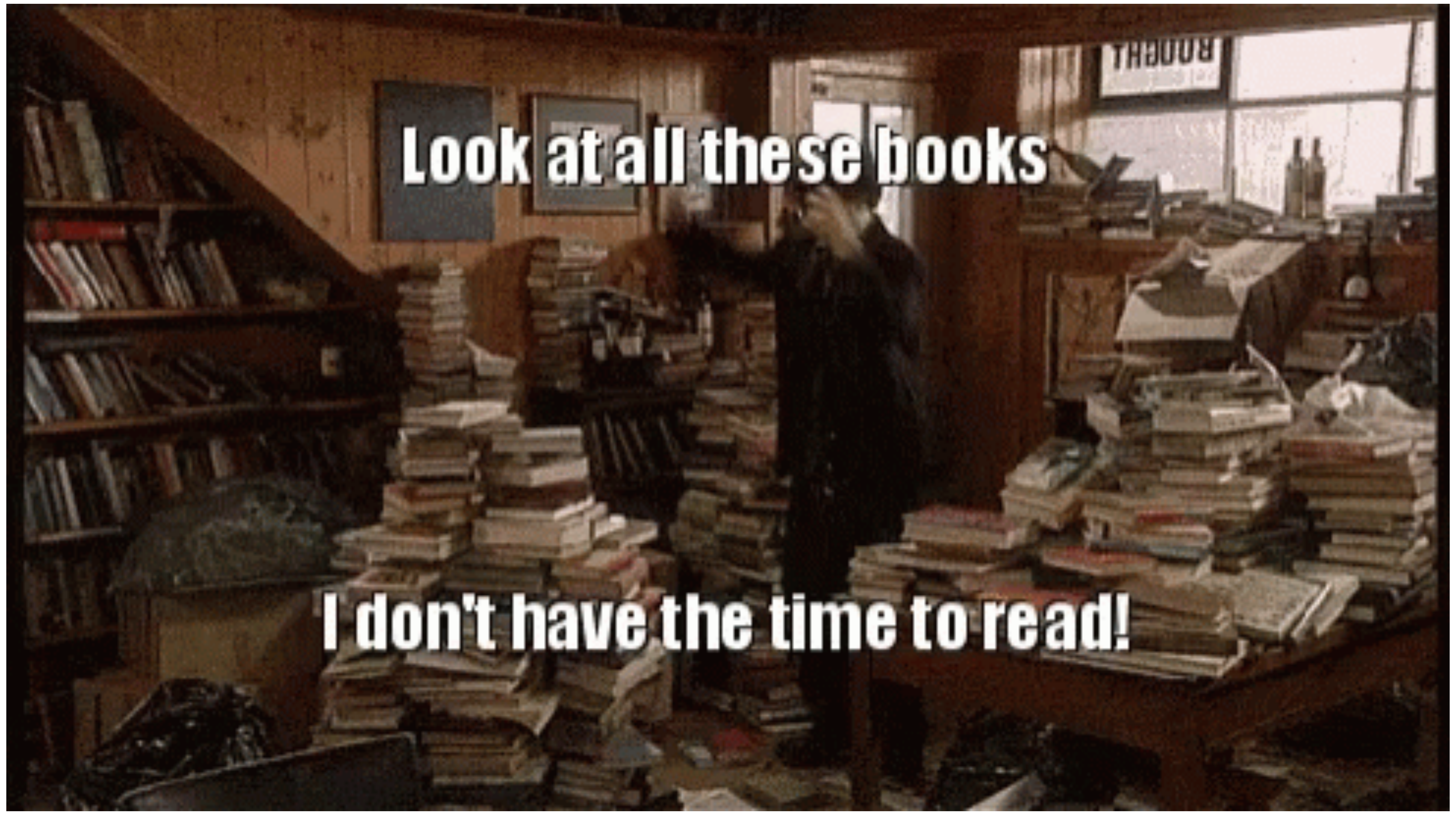
Then go to github + arXiv search

- You might want to apply heuristics to decide which papers you should read first
- But that is later, now let’s assume you have some time…

If both Kaggle & PapersWithCode fail

Then go to github + arXiv search

- You might want to apply heuristics to decide which papers you should read first
- But that is later, now let's assume you have some time…



You are new to the field

and would like to get familiar with it

You are new to the field

and would like to get familiar with it

You are new to the field

and would like to get familiar with it

- The best idea is to start with some modern course (e.g. fast.ai)

You are new to the field

and would like to get familiar with it

- The best idea is to start with some modern course (e.g. fast.ai)
- Starting with dataset to play with (see Kaggle, PapersWithCode) is also a good idea.

You are new to the field

and would like to get familiar with it

- The best idea is to start with some modern course (e.g. fast.ai)
- Starting with dataset to play with (see Kaggle, PapersWithCode) is also a good idea.
- Also a good idea is to start with a good benchmark paper

You are new to the field
and would like to get familiar with it

- The best idea is to start with some modern course (e.g. fast.ai)
- Starting with dataset to play with (see Kaggle, PapersWithCode) is also a good idea.
- Also a good idea is to start with a good benchmark paper
- Or a good survey paper

You are new to the field

and would like to get familiar with it

- The best idea is to start with some modern course (e.g. fast.ai)
- Starting with dataset to play with (see Kaggle, PapersWithCode) is also a good idea.
- Also a good idea is to start with a good benchmark paper
- Or a good survey paper
- But it is not always easy to find good survey paper, there many bad ones

How to spot good benchmark paper?

Citations

- It is used by community (unless it is just published 2 days ago), i.e. people cite it

How to spot good benchmark paper?

Citations

- It is used by community (unless it is just published 2 days ago), i.e. people cite it

https://arxiv.org › cs ▾ Přeložit tuto stránku

Revisiting Oxford and Paris: Large-Scale Image Retrieval ...

autor: F Radenović · 2018 · Počet citací tohoto článku: 203 — In this paper we address issues with image **retrieval** benchmarking on standard and popular **Oxford** 5k and **Paris** 6k datasets...

How to spot good benchmark paper?

Citations

- It is used by community (unless it is just published 2 days ago), i.e. people cite it

https://arxiv.org › cs ▾ Přeložit tuto stránku

Revisiting Oxford and Paris: Large-Scale Image Retrieval ...

autor: F Radenović · 2018 · Počet citací tohoto článku: 203 — In this paper we address issues with image **retrieval** benchmarking on standard and popular **Oxford** 5k and **Paris** 6k datasets...

https://research.google › pub49052 ▾ Přeložit tuto stránku

Google Landmarks Dataset v2 - A Large ... - Google Research

autor: T Weyand · 2020 · Počet citací tohoto článku: 72 — We introduce the **Google Landmarks Dataset** v2 (GLDv2), a new benchmark for large-scale, fine-grained instance recognition and...

How to spot good benchmark paper?

Citations

- It is used by community (unless it is just published 2 days ago), i.e. people cite it

https://arxiv.org › cs ▾ Přeložit tuto stránku

Revisiting Oxford and Paris: Large-Scale Image Retrieval ...

autor: F Radenović · 2018 · Počet citací tohoto článku: 203 — In this paper we address issues with image **retrieval** benchmarking on standard and popular **Oxford** 5k and **Paris** 6k datasets...

https://research.google › pub49052 ▾ Přeložit tuto stránku

Google Landmarks Dataset v2 - A Large ... - Google Research

autor: T Weyand · 2020 · Počet citací tohoto článku: 72 — We introduce the **Google Landmarks Dataset** v2 (GLDv2), a new benchmark for large-scale, fine-grained instance recognition and...

https://arxiv.org › cs ▾ Přeložit tuto stránku

Image Matching across Wide Baselines: From Paper to Practice

autor: Y Jin · 2020 · Počet citací tohoto článku: 78 — Abstract: We introduce a comprehensive benchmark for local features and robust estimation algorithms, focusing on the downstream ta...

Cite as: arXiv:2003.01587

How to spot good benchmark paper?

Citations

- It is used by community (unless it is just published 2 days ago), i.e. people cite it

https://arxiv.org › cs ▾ Přeložit tuto stránku

Revisiting Oxford and Paris: Large-Scale Image Retrieval ...

autor: F Radenović · 2018 · Počet citací tohoto článku: 203 — In this paper we address issues with image **retrieval** benchmarking on standard and popular **Oxford** 5k and **Paris** 6k datasets...

https://research.google › pub49052 ▾ Přeložit tuto stránku

Google Landmarks Dataset v2 - A Large ... - Google Research

autor: T Weyand · 2020 · Počet citací tohoto článku: 72 — We introduce the **Google Landmarks Dataset** v2 (GLDv2), a new benchmark for large-scale, fine-grained instance recognition and...

https://arxiv.org › cs ▾ Přeložit tuto stránku

Image Matching across Wide Baselines: From Paper to Practice

autor: Y Jin · 2020 · Počet citací tohoto článku: 78 — Abstract: We introduce a comprehensive benchmark for local features and robust estimation algorithms, focusing on the downstream ta...

Cite as: arXiv:2003.01587

https://www.robots.ox.ac.uk › vibes_ijcv2004 ▾ PDF

A Comparison of Affine Region Detectors - University of Oxford

autor: K MIKOLAJCZYK · Počet citací tohoto článku: 4188 — Abstract. The paper gives a snapshot of the state of the art in **affine** covariant region **detectors**, and compares their...

How to spot good benchmark paper?

Citations

- It is used by community (unless it is just published 2 days ago), i.e. people cite it

https://arxiv.org › cs ▾ Přeložit tuto stránku

Revisiting Oxford and Paris: Large-Scale Image Retrieval ...

autor: F Radenović · 2018 · Počet citací tohoto článku: 203 — In this paper we address issues with image **retrieval** benchmarking on standard and popular **Oxford** 5k and **Paris** 6k datasets...

https://research.google › pub49052 ▾ Přeložit tuto stránku

Google Landmarks Dataset v2 - A Large ... - Google Research

autor: T Weyand · 2020 · Počet citací tohoto článku: 72 — We introduce the **Google Landmarks Dataset** v2 (GLDv2), a new benchmark for large-scale, fine-grained instance recognition and...

https://arxiv.org › cs ▾ Přeložit tuto stránku

Image Matching across Wide Baselines: From Paper to Practice

autor: Y Jin · 2020 · Počet citací tohoto článku: 78 — Abstract: We introduce a comprehensive benchmark for local features and robust estimation algorithms, focusing on the downstream ta...

Cite as: arXiv:2003.01587

https://www.robots.ox.ac.uk › vibes_ijcv2004 ▾ PDF

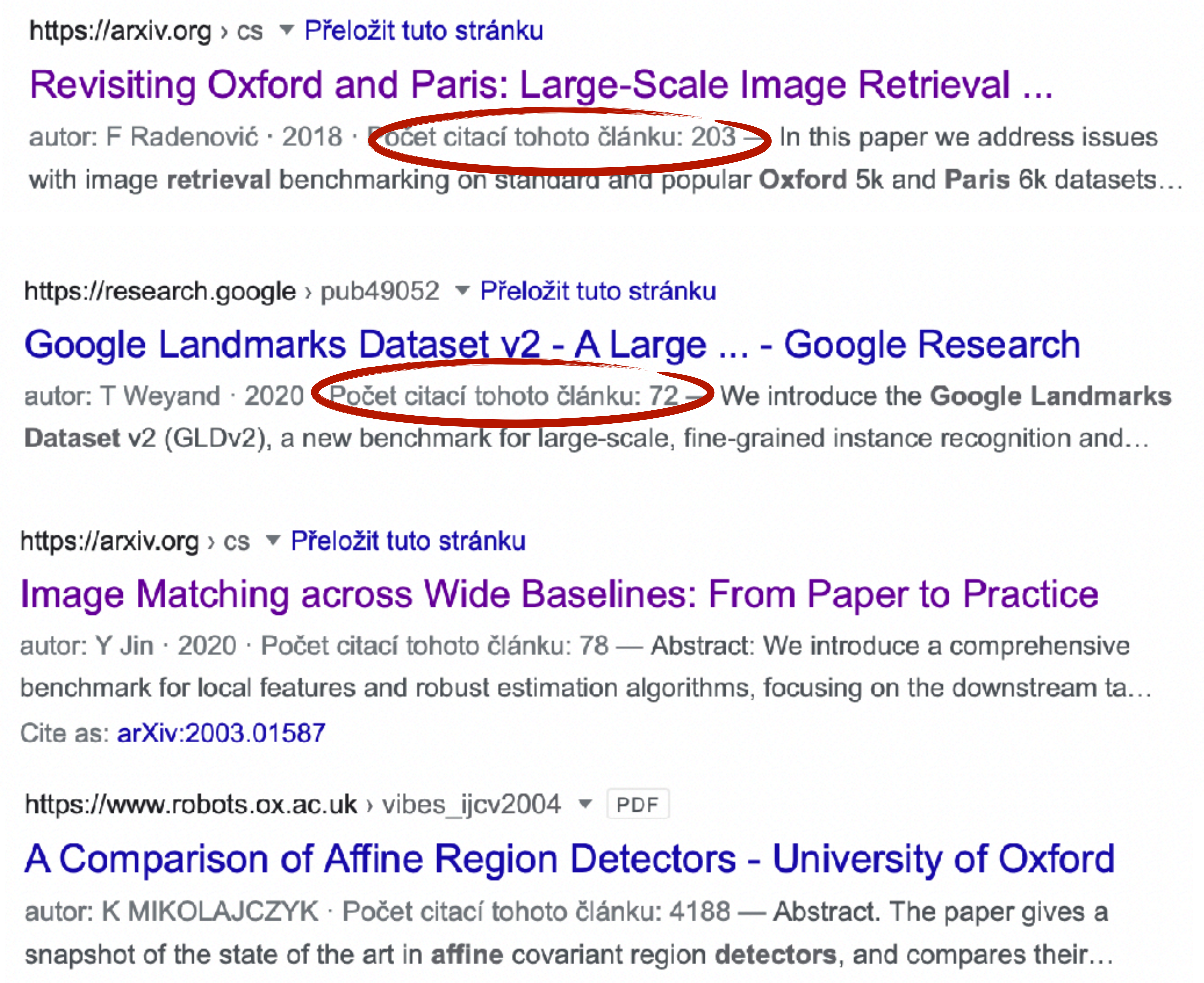
A Comparison of Affine Region Detectors - University of Oxford

autor: K MIKOLAJCZYK · Počet citací tohoto článku: 4188 — Abstract. The paper gives a snapshot of the state of the art in **affine** covariant region **detectors**, and compares their...

How to spot good benchmark paper?

Citations

- It is used by community (unless it is just published 2 days ago), i.e. people cite it



https://arxiv.org › cs ▾ Přeložit tuto stránku

Revisiting Oxford and Paris: Large-Scale Image Retrieval ...

autor: F Radenović · 2018 · Počet citací tohoto článku: 203 — In this paper we address issues with image **retrieval** benchmarking on standard and popular **Oxford** 5k and **Paris** 6k datasets...

https://research.google › pub49052 ▾ Přeložit tuto stránku

Google Landmarks Dataset v2 - A Large ... - Google Research

autor: T Weyand · 2020 · Počet citací tohoto článku: 72 — We introduce the **Google Landmarks Dataset** v2 (GLDv2), a new benchmark for large-scale, fine-grained instance recognition and...

https://arxiv.org › cs ▾ Přeložit tuto stránku

Image Matching across Wide Baselines: From Paper to Practice

autor: Y Jin · 2020 · Počet citací tohoto článku: 78 — Abstract: We introduce a comprehensive benchmark for local features and robust estimation algorithms, focusing on the downstream ta...

Cite as: arXiv:2003.01587

https://www.robots.ox.ac.uk › vibes_ijcv2004 ▾ PDF

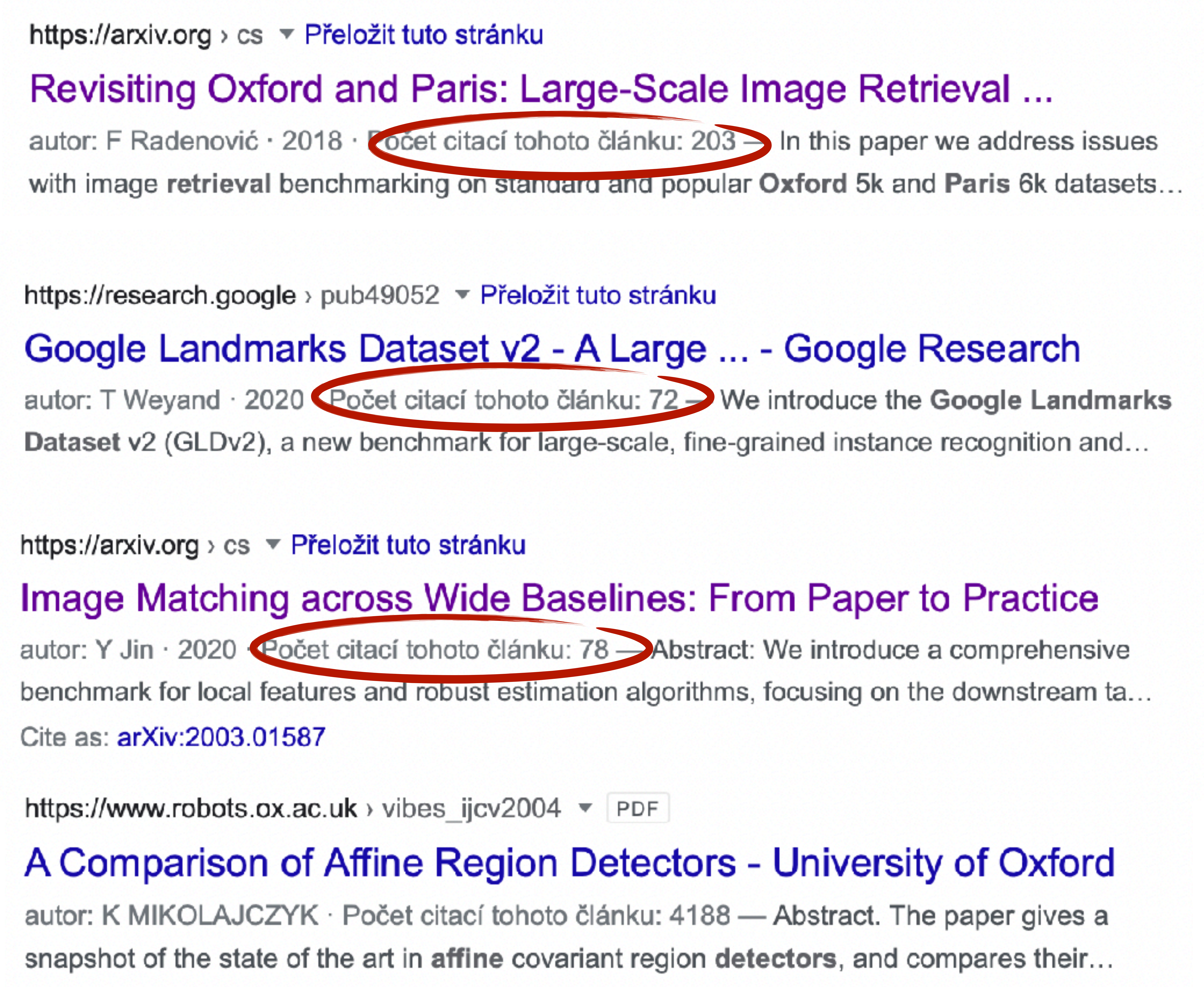
A Comparison of Affine Region Detectors - University of Oxford

autor: K MIKOLAJCZYK · Počet citací tohoto článku: 4188 — Abstract. The paper gives a snapshot of the state of the art in **affine** covariant region **detectors**, and compares their...

How to spot good benchmark paper?

Citations

- It is used by community (unless it is just published 2 days ago), i.e. people cite it



How to spot good benchmark paper?

Citations

- It is used by community (unless it is just published 2 days ago), i.e. people cite it

https://arxiv.org › cs ▾ Přeložit tuto stránku

Revisiting Oxford and Paris: Large-Scale Image Retrieval ...

autor: F Radenović · 2018 · Počet citací tohoto článku: 203 — In this paper we address issues with image **retrieval** benchmarking on standard and popular **Oxford** 5k and **Paris** 6k datasets...

https://research.google › pub49052 ▾ Přeložit tuto stránku

Google Landmarks Dataset v2 - A Large ... - Google Research

autor: T Weyand · 2020 · Počet citací tohoto článku: 72 — We introduce the **Google Landmarks Dataset** v2 (GLDv2), a new benchmark for large-scale, fine-grained instance recognition and...

https://arxiv.org › cs ▾ Přeložit tuto stránku

Image Matching across Wide Baselines: From Paper to Practice

autor: Y Jin · 2020 · Počet citací tohoto článku: 78 — Abstract: We introduce a comprehensive benchmark for local features and robust estimation algorithms, focusing on the downstream ta...

Cite as: arXiv:2003.01587

https://www.robots.ox.ac.uk › vibes_ijcv2004 ▾ PDF

A Comparison of Affine Region Detectors - University of Oxford

autor: K MIKOLAJCZYK · Počet citací tohoto článku: 4188 — Abstract. The paper gives a snapshot of the state of the art in **affine** covariant region **detectors**, and compares their...

How to spot good benchmark paper?

Year

How to spot good benchmark paper?

Year

- It should not be too old: benchmark from 2004 is a little bit obsolete
- Oxford-Affine is GREAT paper. But it would not help you find modern local features

How to spot good benchmark paper?

Year

- It should not be too old: benchmark from 2004 is a little bit obsolete
- Oxford-Affine is GREAT paper. But it would not help you find modern local features

https://www.robots.ox.ac.uk › vibes_ijcv2004 ▾ PDF

A Comparison of Affine Region Detectors - University of Oxford

autor: K MIKOLAJCZYK · Počet citací tohoto článku: 4188 — Abstract. The paper gives a snapshot of the state of the art in **affine** covariant region **detectors**, and compares their…

How to spot good benchmark paper?

Year

- It should not be too old: benchmark from 2004 is a little bit obsolete

- Oxford-Affine is GREAT paper. But it would not help you find modern local features

How to spot good benchmark paper?

It proposes new dataset (or new annotation)

How to spot good benchmark paper?

It proposes new dataset (or new annotation)

- There are TONS of papers, which just run methods from OpenCV on existing (and old) datasets

How to spot good benchmark paper?

It proposes new dataset (or new annotation)

- There are TONS of papers, which just run methods from OpenCV on existing (and old) datasets

International Journal of
Applied Mathematics,
Electronics and Computers

IJAMEC

Advanced Technology and Science

ISSN: 2147-8228 www.atscience.org/IJAMEC Original Research Paper

A Comparative Evaluation of Well-known Feature Detectors and Descriptors#

How to spot good benchmark paper?

It proposes new dataset (or new annotation)

- There are TONS of papers, which just run methods from OpenCV on existing (and old) datasets

International Journal of
Applied Mathematics,
Electronics and Computers

IJAMEC

ISSN: 2147-8228 www.atscience.org/IJAMEC

Science
Advanced Technology and Science

Original Research Paper

A Comparative Evaluation of Well-known Feature Detectors and Descriptors#

3.7. Effects of JPEG Compression

In this experiment, impacts of compression is examined in terms of comparison of results obtained from each method conducting on ubc dataset. For this purpose, the artifacts have introduced in Joint Photographic Experts Group (JPEG) compression by using a standard xv image browser with the image quality parameter changing from 40% to 2%.

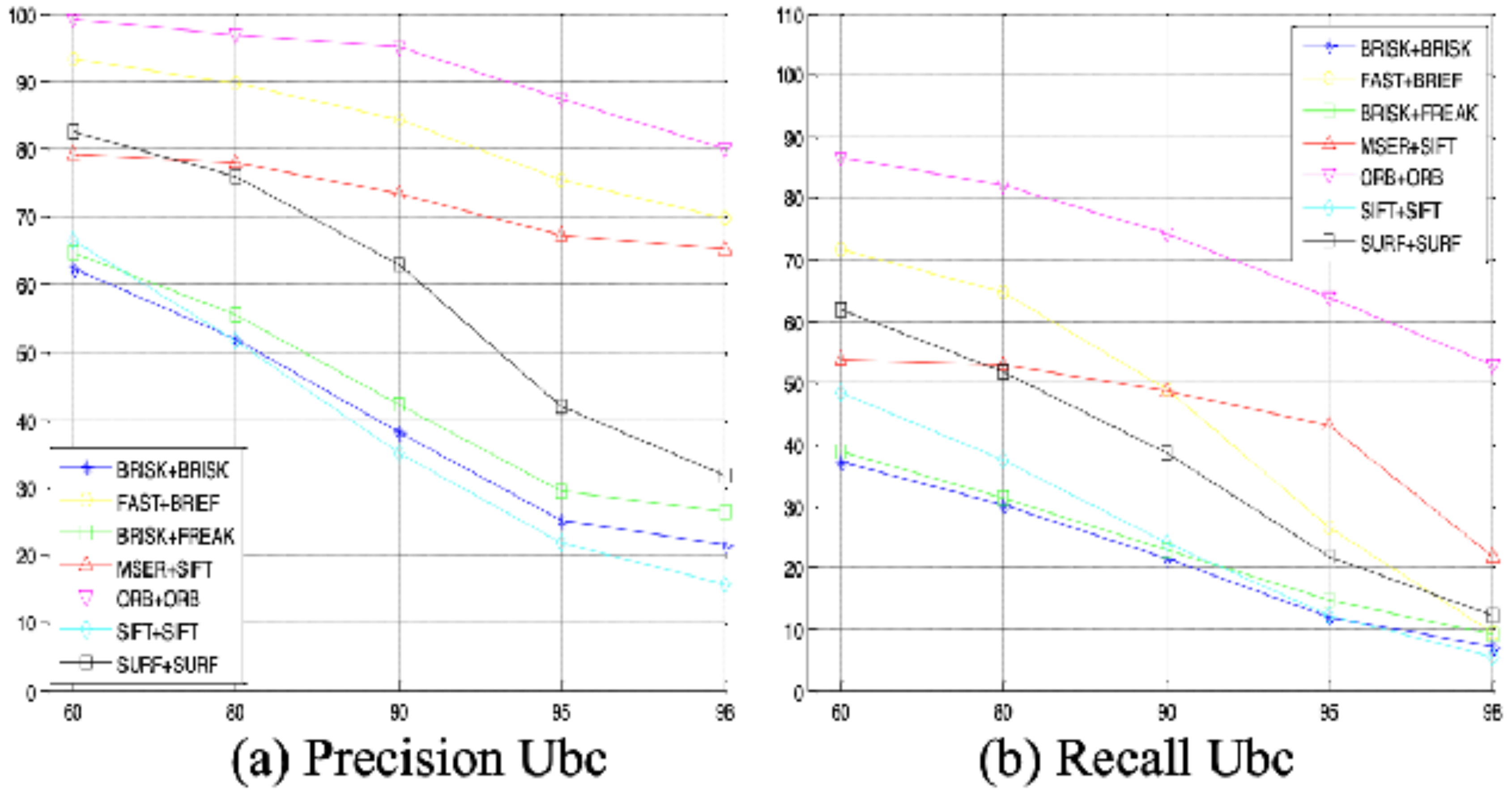
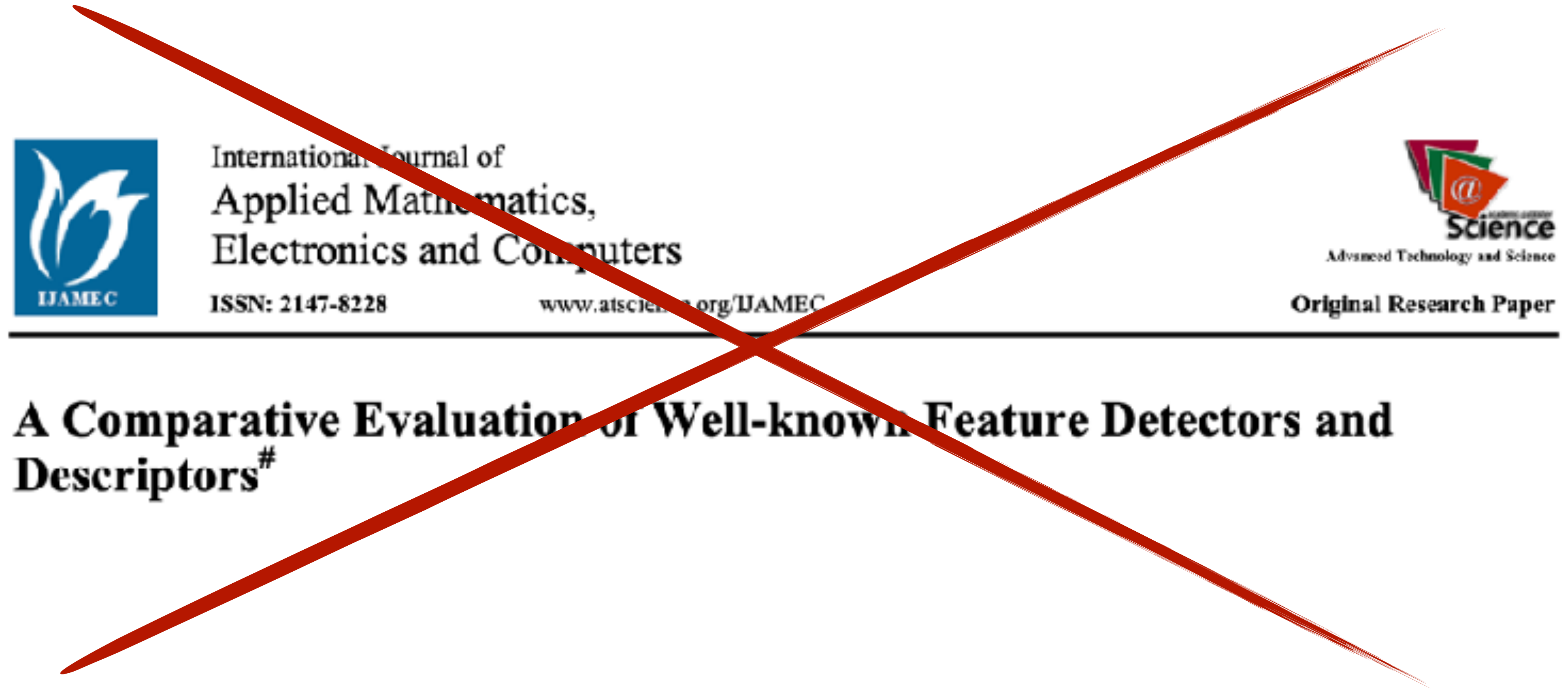


Figure 7: Precision and recall values for JPEG compression.

How to spot good benchmark paper?

It proposes new dataset (or new annotation)

- There are TONS of papers, which just run methods from OpenCV on existing (and old) datasets



International Journal of
Applied Mathematics, Electronics and Computers

ISSN: 2147-8228 www.atscience.org/IJAMEC

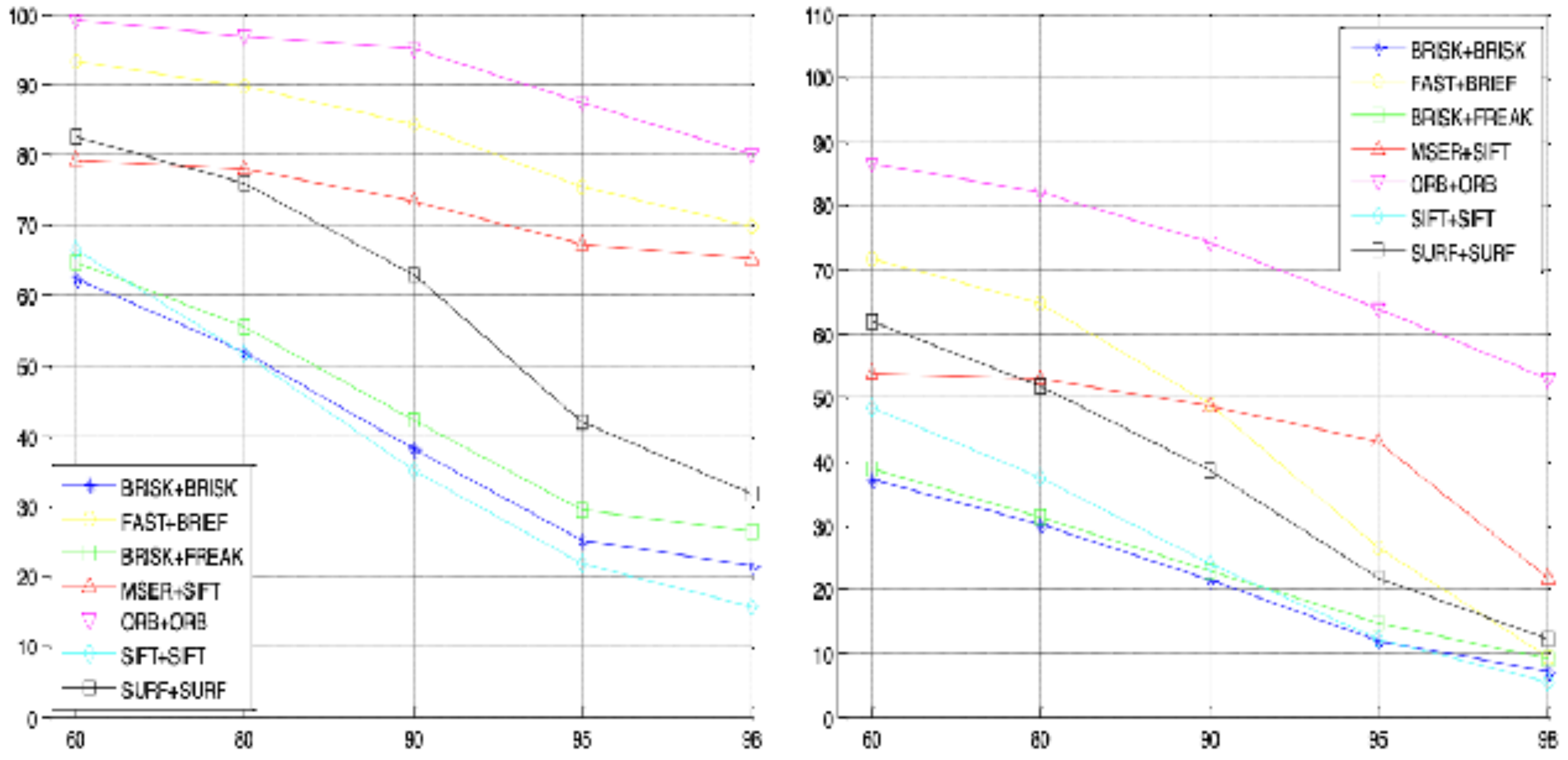
Advanced Technology and Science

Original Research Paper

A Comparative Evaluation of Well-known Feature Detectors and Descriptors#

3.7. Effects of JPEG Compression

In this experiment, impacts of compression is examined in terms of comparison of results obtained from each method conducting on ubc dataset. For this purpose, the artifacts have introduced in Joint Photographic Experts Group (JPEG) compression by using a standard xv image browser with the image quality parameter changing from 40% to 2%.



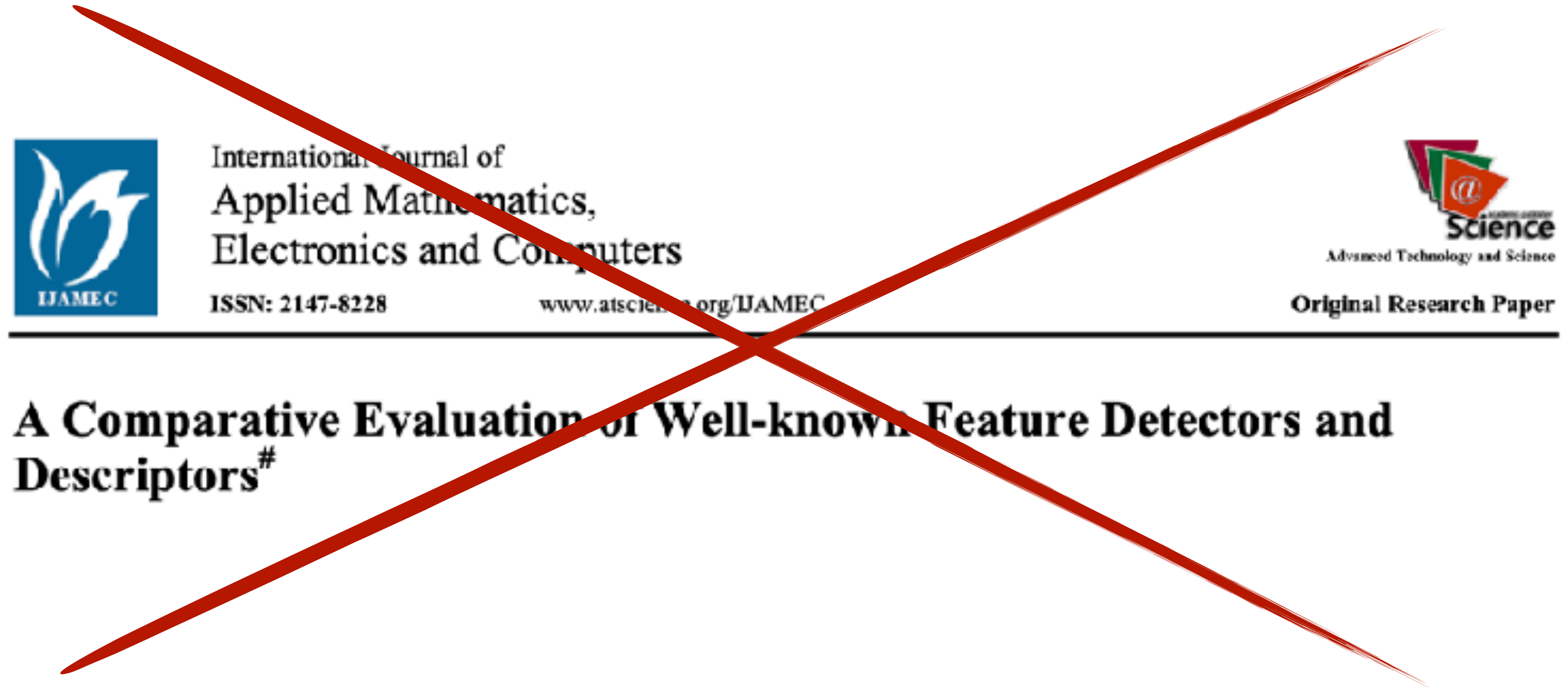
(a) Precision Ubc (b) Recall Ubc

Figure 7: Precision and recall values for JPEG compression.

How to spot good benchmark paper?

It proposes new dataset (or new annotation)

- There are TONS of papers, which just run methods from OpenCV on existing (and old) datasets



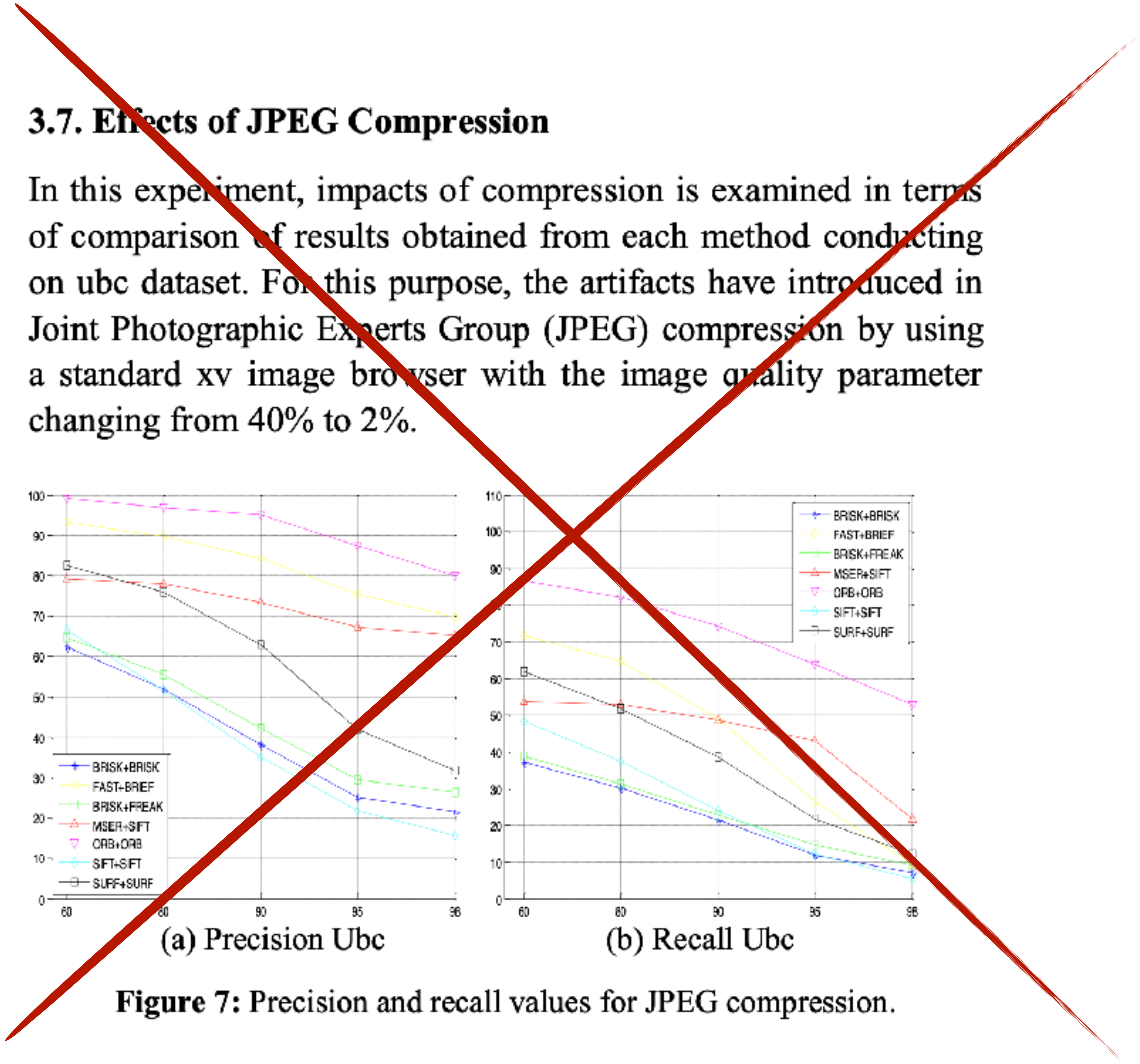
International Journal of
Applied Mathematics,
Electronics and Computers

ISSN: 2147-8228 www.atscience.org/IJAMEC

Advanced Technology and Science

Original Research Paper

A Comparative Evaluation of Well-known Feature Detectors and Descriptors#



3.7. Effects of JPEG Compression

In this experiment, impacts of compression is examined in terms of comparison of results obtained from each method conducting on ubc dataset. For this purpose, the artifacts have introduced in Joint Photographic Experts Group (JPEG) compression by using a standard xv image browser with the image quality parameter changing from 40% to 2%.

(a) Precision Ubc (b) Recall Ubc

Figure 7: Precision and recall values for JPEG compression.

How to spot good benchmark paper?

It has a challenge linked to it, often — many years

https://www.visuallocalization.net/benchmark/

https://www.votchallenge.net

How to spot good benchmark paper?

It has a challenge linked to it, often — many years

LONG-TERM VISUAL LOCALIZATION

HOME DATASETS BENCHMARK SUBMISSION ECCV 2020 WORKSHOP

Benchmark

Listed below are the public results on the three benchmark datasets. The localization results are reported as the perc
translation and rotation thresholds, for each condition.

Aachen Day-Night dataset

Localization thresholds:

All conditions: (0.25m, 2°) / (0.5m, 5°) / (5m, 10°)

Method	day
Combined tuned HLoc	89.7 / 95.8 / 98
Hierarchical Localization - SuperPoint + SuperGlue	89.6 / 95.4 / 98
SuperGlue + Patch2Pix (HLoc)	89.2 / 95.5 / 98
ONavi-V2.0	88.6 / 95.1 / 99
hloc	89.3 / 95.9 / 98
KAPTURE-R2D2-FUSION	89.4 / 96.4 / 99

https://www.visuallocalization.net/benchmark/

https://www.votchallenge.net

How to spot good benchmark paper?

It has a challenge linked to it, often — many years

LONG-TERM VISUAL LOCALIZATION

HOME DATASETS BENCHMARK SUBMISSION ECCV 2020 WORKSHOP

Benchmark

Listed below are the public results on the three benchmark datasets. The localization results are reported as the perc
translation and rotation thresholds, for each condition.

Aachen Day-Night dataset

Localization thresholds:

All conditions: (0.25m, 2°) / (0.5m, 5°) / (5m, 10°)

Method	day
Combined tuned HLoc	89.7 / 95.8 / 98
Hierarchical Localization - SuperPoint + SuperGlue	89.6 / 95.4 / 98
SuperGlue + Patch2Pix (HLoc)	89.2 / 95.5 / 98
ONavi-V2.0	88.6 / 95.1 / 99
hloc	89.3 / 95.9 / 98
KAPTURE-R2D2-FUSION	89.4 / 96.4 / 99

https://www.visuallocalization.net/benchmark/

Benchmarks

Official results of past challenges are still available as a benchmark to researchers. We invite the visual tracking researchers to compare their methods with those results.

VOT2021 benchmark

The VOT2021 benchmark addresses short-term, long-term, real-time, RGB and RGBD trackers. Results were presented at the ICCV2021 VOT workshop.

VOT2020 benchmark

The VOT2020 benchmark addresses short-term, long-term, real-time, RGB, RGBT and RGBD trackers. Results were presented at the ECCV2020 VOT workshop.

VOT2019 benchmark

The VOT2019 benchmark addresses short-term, long-term, real-time, RGB, RGBT and RGBD trackers. Results were presented at the ICCV2019 VOT workshop.

VOT2018 benchmark

The VOT2018 benchmark introduced a long-term subchallenge VOT-LT2018. Results were presented at the VOT workshop at ECCV2018.

VOT2017 benchmark

The VOT2017 benchmark introduced a refreshed a dataset and a real-time experiment. The winner will was determined on sequestered dataset. The results were presented at the VOT workshop at ICCV2017.

VOT2016 benchmark

The fourth challenge updated the dataset of 60 sequences with new annotations. The results were published in a joint paper presented at a workshop at ECCV2016.

https://www.votchallenge.net

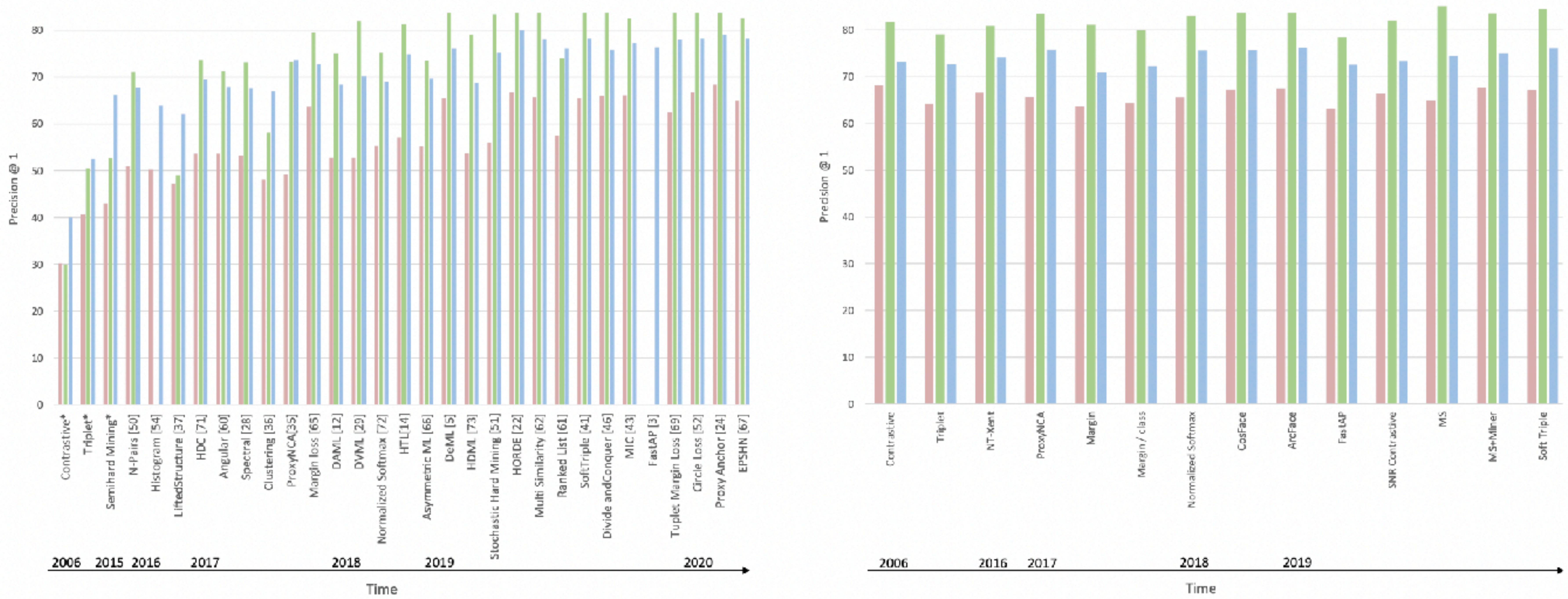
How to spot good benchmark paper?

It re-evaluates existing method on the common ground

https://arxiv.org › cs ▾ Přeložit tuto stránku

[2003.08505] A Metric Learning Reality Check - arXiv

autor: K Musgrave · 2020 · Počet citací tohoto článku: 159 — Abstract: Deep **metric learning** papers from the past four years have consistently claimed great advances in accuracy, often...



(a) The trend according to papers (b) The trend according to reality

How to spot good benchmark paper?

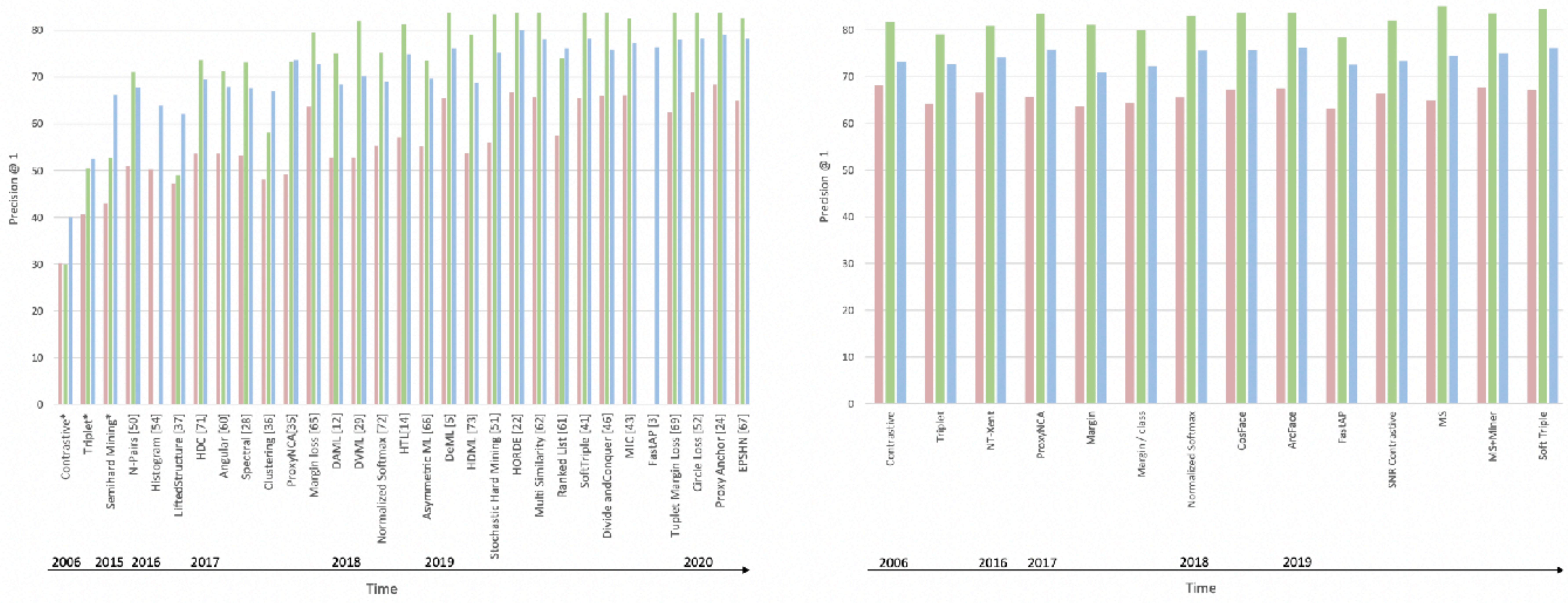
It re-evaluates existing method on the common ground

- Some papers do not propose new dataset, but they critically evaluate existing metrics

https://arxiv.org › cs ▾ Přeložit tuto stránku

[2003.08505] A Metric Learning Reality Check - arXiv

autor: K Musgrave · 2020 · Počet citací tohoto článku: 159 — Abstract: Deep **metric learning** papers from the past four years have consistently claimed great advances in accuracy, often...



(a) The trend according to papers (b) The trend according to reality

How to spot good benchmark paper?

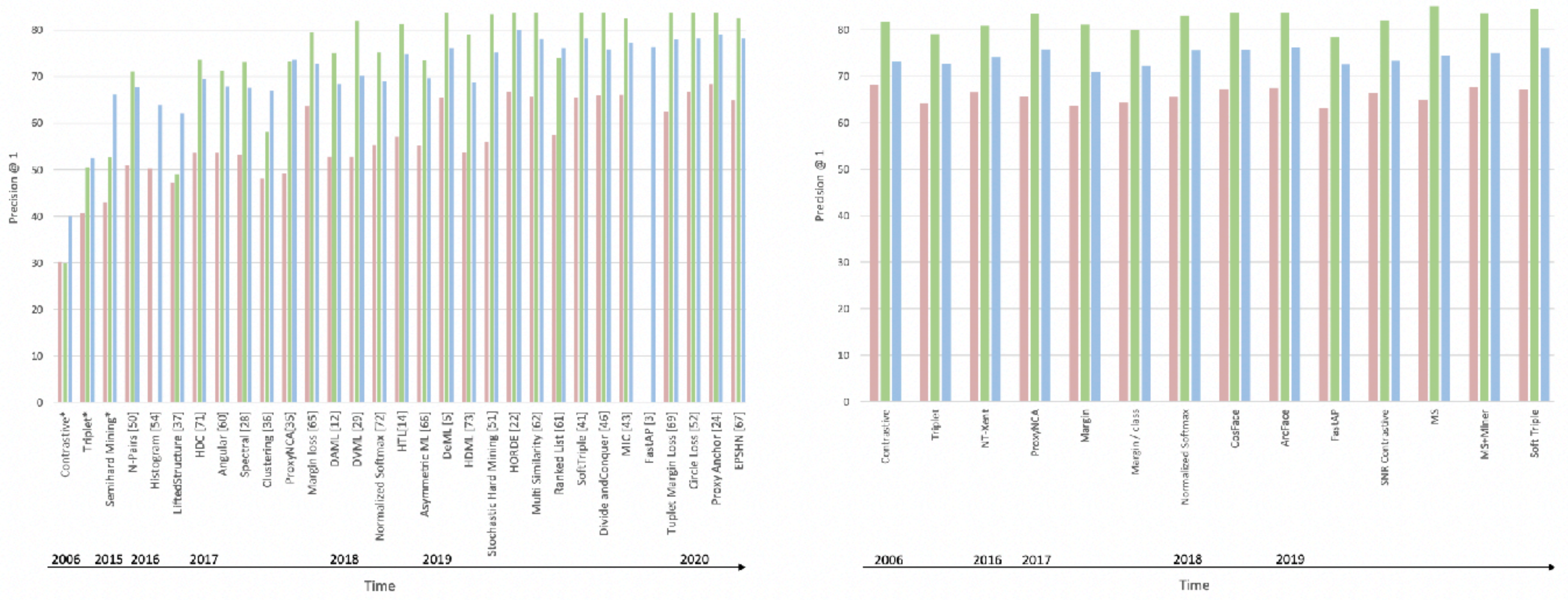
It re-evaluates existing method on the common ground

- Some papers do not propose new dataset, but they critically evaluate existing metrics
- And/or re-implement the methods from scratch on the common codebase

https://arxiv.org › cs ▾ Přeložit tuto stránku

[2003.08505] A Metric Learning Reality Check - arXiv

autor: K Musgrave · 2020 · Počet citací tohoto článku: 159 — Abstract: Deep **metric learning** papers from the past four years have consistently claimed great advances in accuracy, often...



(a) The trend according to papers (b) The trend according to reality

Good Surveys: harder to distinguish

But there are couple of heuristics

Good Surveys: harder to distinguish

But there are couple of heuristics

- Good modern survey is getting updated

Good Surveys: harder to distinguish

But there are couple of heuristics

- Good modern survey is getting updated

arXiv:1704.05519 (cs)

[Submitted on 18 Apr 2017 (v1), last revised 17 Mar 2021 (this version, v3)]

Computer Vision for Autonomous Vehicles: Problems, Datasets and State of the Art

Joel Janai, Fatma Güney, Aseem Behl, Andreas Geiger

Good Surveys: harder to distinguish

But there are couple of heuristics

- Good modern survey is getting updated

arXiv:1704.05519 (cs)

[Submitted on 18 Apr 2017 (v1), last revised 17 Mar 2021 (this version, v3)]

Computer Vision for Autonomous Vehicles: Problems, Datasets and State of the Art

Joel Janai, Fatma Güney, Aseem Behl, Andreas Geiger

Good Surveys: harder to distinguish

But there are couple of heuristics

- Good modern survey is getting updated

arXiv:1704.05519 (cs)

[Submitted on 18 Apr 2017 (v1), last revised 17 Mar 2021 (this version, v3)]

Computer Vision for Autonomous Vehicles: Problems, Datasets and State of the Art

Joel Janai, Fatma Güney, Aseem Behl, Andreas Geiger

[Submitted on 26 Jul 2018 (v1), last revised 14 Jun 2019 (this version, v3)]

From handcrafted to deep local features

Gabriela Csurka, Christopher R. Dance, Martin Humenberger

Good Surveys: harder to distinguish

But there are couple of heuristics

- Good modern survey is getting updated

arXiv:1704.05519 (cs)

[Submitted on 18 Apr 2017 (v1), last revised 17 Mar 2021 (this version, v3)]

Computer Vision for Autonomous Vehicles: Problems, Datasets and State of the Art

Joel Janai, Fatma Güney, Aseem Behl, Andreas Geiger

[Submitted on 26 Jul 2018 (v1), last revised 14 Jun 2019 (this version, v3)]

From handcrafted to deep local features

Gabriela Csurka, Christopher R. Dance, Martin Humenberger

Good Surveys: harder to distinguish

But there are couple of heuristics

- Good modern survey is getting updated

- available on arXiv/author page

arXiv:1704.05519 (cs)

[Submitted on 18 Apr 2017 (v1), last revised 17 Mar 2021 (this version, v3)]

Computer Vision for Autonomous Vehicles: Problems, Datasets and State of the Art

Joel Janai, Fatma Güney, Aseem Behl, Andreas Geiger

[Submitted on 26 Jul 2018 (v1), last revised 14 Jun 2019 (this version, v3)]

From handcrafted to deep local features

Gabriela Csurka, Christopher R. Dance, Martin Humenberger

Good Surveys: harder to distinguish

But there are couple of heuristics

- Good modern survey is getting updated
- available on arXiv/author page

arXiv:1704.05519 (cs)

[Submitted on 18 Apr 2017 (v1), last revised 17 Mar 2021 (this version, v3)]

Computer Vision for Autonomous Vehicles: Problems, Datasets and State of the Art

Joel Janai, Fatma Güney, Aseem Behl, Andreas Geiger

[Submitted on 26 Jul 2018 (v1), last revised 14 Jun 2019 (this version, v3)]

From handcrafted to deep local features

Gabriela Csurka, Christopher R. Dance, Martin Humenberger

[Submitted on 14 Sep 2020 (v1), last revised 16 Sep 2020 (this version, v2)]

Efficient Transformers: A Survey

Yi Tay, Mostafa Dehghani, Dara Bahri, Donald Metzler

Good Surveys: harder to distinguish

But there are couple of heuristics

- Good modern survey is getting updated

- available on arXiv/author page

- bad survey is “published" in obscure journals or behind the paywall

arXiv:1704.05519 (cs)

[Submitted on 18 Apr 2017 (v1), last revised 17 Mar 2021 (this version, v3)]

Computer Vision for Autonomous Vehicles: Problems, Datasets and State of the Art

Joel Janai, Fatma Güney, Aseem Behl, Andreas Geiger

[Submitted on 26 Jul 2018 (v1), last revised 14 Jun 2019 (this version, v3)]

From handcrafted to deep local features

Gabriela Csurka, Christopher R. Dance, Martin Humenberger

[Submitted on 14 Sep 2020 (v1), last revised 16 Sep 2020 (this version, v2)]

Efficient Transformers: A Survey

Yi Tay, Mostafa Dehghani, Dara Bahri, Donald Metzler

Good Surveys: harder to distinguish

But there are couple of heuristics

- Good modern survey is getting updated

- available on arXiv/author page

- bad survey is "published" in obscure journals or behind the paywall

http://www.ijeetc.com › 2_NCETEC001_(p.9-17).... ▾ PDF

A SURVEY OF TRENDS IN LOCAL INVARIANT FEATURE ...

autor: JN Surekha · 2017 · Počet citací tohoto článku: 2 — Many computer vision algorithms use **feature detection** as the initial step, so as a result, a very large number of **feature detectors**...

9 stránek

arXiv:1704.05519 (cs)

[Submitted on 18 Apr 2017 (v1), last revised 17 Mar 2021 (this version, v3)]

Computer Vision for Autonomous Vehicles: Problems, Datasets and State of the Art

Joel Janai, Fatma Güney, Aseem Behl, Andreas Geiger

[Submitted on 26 Jul 2018 (v1), last revised 14 Jun 2019 (this version, v3)]

From handcrafted to deep local features

Gabriela Csurka, Christopher R. Dance, Martin Humenberger

[Submitted on 14 Sep 2020 (v1), last revised 16 Sep 2020 (this version, v2)]

Efficient Transformers: A Survey

Yi Tay, Mostafa Dehghani, Dara Bahri, Donald Metzler

Good Surveys: harder to distinguish

But there are couple of heuristics

- Good modern survey is getting updated
- available on arXiv/author page
- bad survey is “published" in obscure journals or behind the paywall

~~http://www.ijeetc.com › 2_NCETEC001_(p.9-17).... PDF~~
~~A SURVEY OF TRENDS IN LOCAL INVARIANT FEATURE ...~~
~~autor: JN Surekha · 2017 · Počet citací … článku: 2 — Many computer vision algorithms use **feature detection** as the initial step, so as a result, a very large number of **feature detectors**…~~
~~9 stránek~~

arXiv:1704.05519 (cs)

[Submitted on 18 Apr 2017 (v1), last revised 17 Mar 2021 (this version, v3)]

Computer Vision for Autonomous Vehicles: Problems, Datasets and State of the Art

Joel Janai, Fatma Güney, Aseem Behl, Andreas Geiger

[Submitted on 26 Jul 2018 (v1), last revised 14 Jun 2019 (this version, v3)]

From handcrafted to deep local features

Gabriela Csurka, Christopher R. Dance, Martin Humenberger

[Submitted on 14 Sep 2020 (v1), last revised 16 Sep 2020 (this version, v2)]

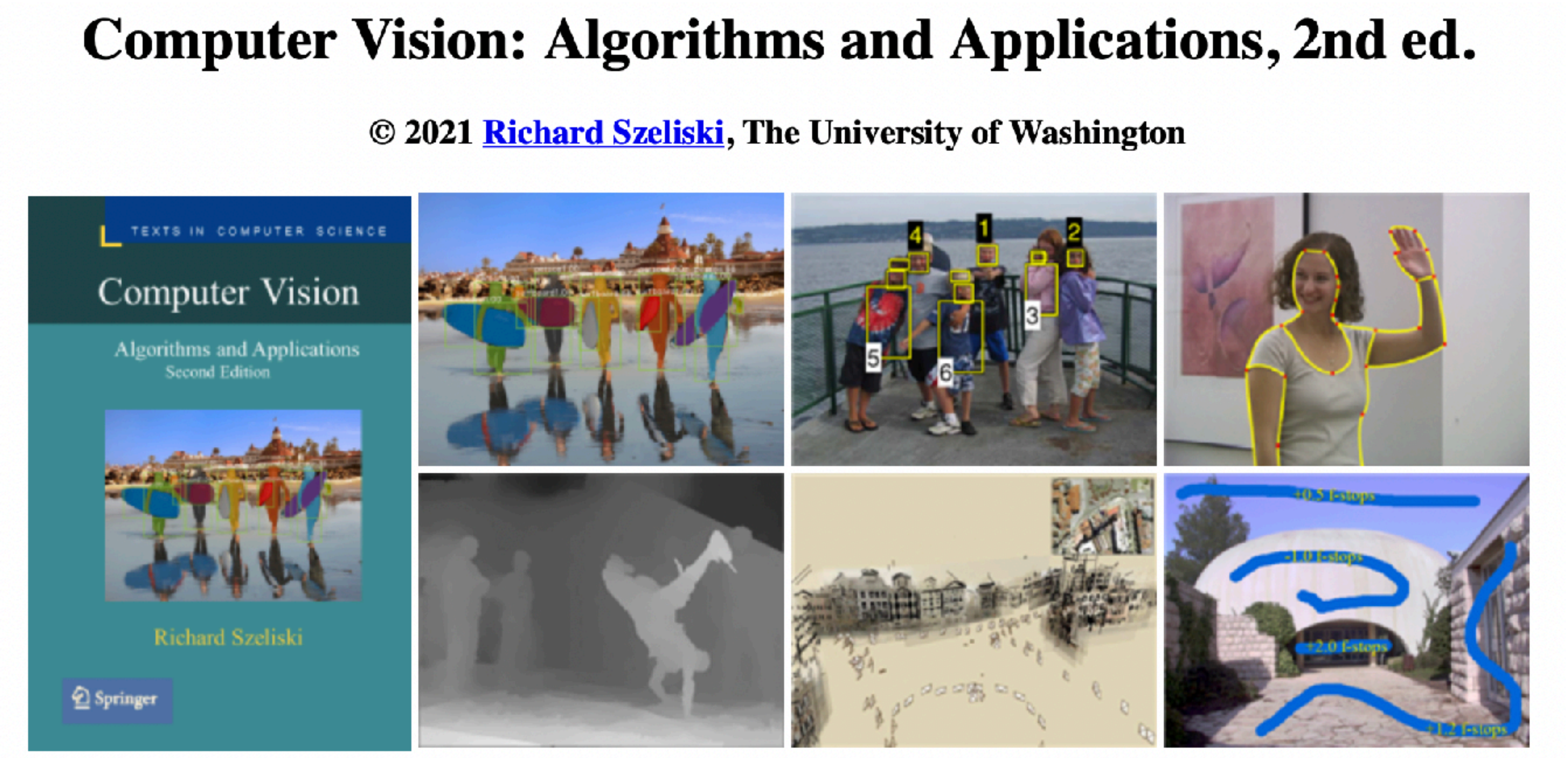
Efficient Transformers: A Survey

Yi Tay, Mostafa Dehghani, Dara Bahri, Donald Metzler

Good Surveys: harder to evaluate

at least without experience

- Cheat: you can use new edition (2021) of Szeliski book as a meta-survey.



https://szeliski.org/Book/

How to read benchmark/survey

How to read benchmark/survey paper

if you are new to the field

How to read benchmark/survey paper

if you are new to the field

- First, understand the task, by reading the intro and thinking. Write down:

How to read benchmark/survey paper

if you are new to the field

- First, understand the task, by reading the intro and thinking. Write down:
- Who are the “users" of the algorithm?

How to read benchmark/survey paper

if you are new to the field

- First, understand the task, by reading the intro and thinking. Write down:
- Who are the “users" of the algorithm?
    - Where is the result is going to be used?

How to read benchmark/survey paper

if you are new to the field

- First, understand the task, by reading the intro and thinking. Write down:
- Who are the “users" of the algorithm?
    - Where is the result is going to be used?
    - What is important for the downstream task?

How to read benchmark/survey paper

if you are new to the field

- First, understand the task, by reading the intro and thinking. Write down:
- Who are the “users" of the algorithm?
  - Where is the result is going to be used?
  - What is important for the downstream task?
  - How one can measure it?

How to read benchmark/survey paper

if you are new to the field

- First, understand the task, by reading the intro and thinking. Write down:
- Who are the “users" of the algorithm?
  - Where is the result is going to be used?
  - What is important for the downstream task?
  - How one can measure it?
  - Where the data comes from in production?

How to read benchmark/survey paper

if you are new to the field

- First, understand the task, by reading the intro and thinking. Write down:
- Who are the “users" of the algorithm?
  - Where is the result is going to be used?
  - What is important for the downstream task?
  - How one can measure it?
  - Where the data comes from in production?
  - Can we avoid solving this task at all?

How to read benchmark/survey paper

if you are new to the field

- First, understand the task, by reading the intro and thinking. Write down:
- Who are the “users" of the algorithm?
  - Where is the result is going to be used?
  - What is important for the downstream task?
  - How one can measure it?
  - Where the data comes from in production?
  - Can we avoid solving this task at all?

Noname manuscript No.
(will be inserted by the editor)

Image Matching Across Wide Baselines: From Paper to Practice

Yuhe Jin · Dmytro Mishkin · Anastasiia Mishchuk · Jiri Matas · Pascal Fua · Kwang Moo Yi · Eduard Trulls

arXiv:2003.01587v5 [cs.CV] 11 Feb 2021

Received: date / Accepted: date

Abstract We introduce a comprehensive benchmark for local features and robust estimation algorithms, focusing on the downstream task – the accuracy of the reconstructed camera pose – as our primary metric. Our pipeline's modular structure allows easy integration, configuration, and combination of different methods and heuristics. This is demonstrated by embedding dozens of popular algorithms and evaluating them, from seminal works to the cutting edge of machine learning research. We show that with proper settings, classical solutions may still outperform the *perceived state of the art*.

Besides establishing the *actual state of the art*, the conducted experiments reveal unexpected properties of Structure from Motion (SfM) pipelines that can help improve their performance, for both algorithmic and learned methods. Data and code are online[1], providing an easy-to-use and flexible framework for the benchmarking of local features and robust estimation methods, both *alongside* and *against* top-performing methods. This work provides a basis for the Image Matching Challenge[2].

Fig. 1 Every paper claims to outperform the state of the art. Is this possible, or an artifact of insufficient validation? On the left, we show stereo matches obtained with **D2-Net** (2019) [37], a state-of-the-art local feature, using OpenCV RANSAC with its default settings. We color the inliers in green if they are correct and in red otherwise. On the right, we show **SIFT** (1999) [54] with a carefully tuned MAGSAC [15] – notice how the latter performs much better. This illustrates our take-home message: to correctly evaluate a method's performance, it needs to be embedded within the pipeline used to solve a given problem, and the different components in said pipeline need to be tuned carefully and jointly, which requires engineering and domain expertise. We fill this need with a new, modular benchmark for sparse image matching, incorporating dozens of built-in methods.

This work was partially supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant “Deep Visual Geometry Machines” (RGPIN-2018-03788), by systems supplied by Compute Canada, and by Google's Visual Positioning Service. DM and JM were supported by OP VVV funded project CZ.02.1.01/0.0/0.0/16 019/0000765 “Research Center for Informatics”. DM was also supported by CTU student grant SGS17/185/OHK3/3T/13 and by the Austrian Ministry for Transport, Innovation and Technology, the Federal Ministry for Digital and Economic Affairs, and the Province of Upper Austria in the frame of the COMET center SCCH. AM was supported by the Swiss National Science Foundation.

Y. Jin, K.M. Yi
Visual Computing Group, University of Victoria
E-mail: {yuhejin, kyi}@uvic.ca

A. Mishchuk, P. Fua
Computer Vision Lab, École Polytechnique Fédérale de Lausanne
E-mail: {anastasiia.mishchuk, pascal.fua}@epfl.ch

D. Mishkin, J. Matas
Center for Machine Perception, Czech Technical University in Prague
E-mail: {mishkdmy, matas}@cmp.felk.cvut.cz

E. Trulls
Google Research
E-mail: trulls@google.com

[1] https://github.com/vcg-uvic/image-matching-benchmark
[2] https://vision.uvic.ca/image-matching-challenge

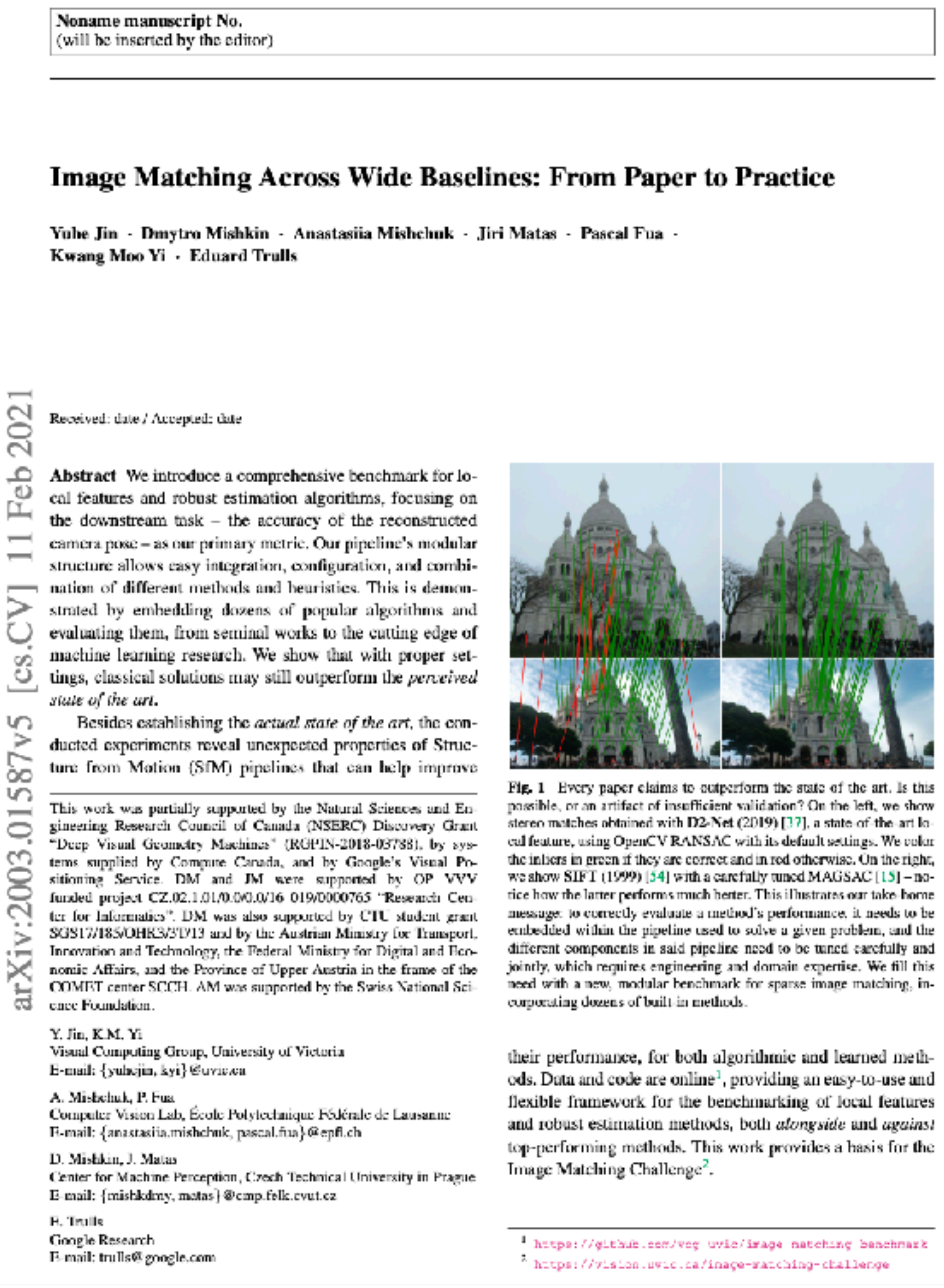
How to read benchmark/survey paper

if you are new to the field

- First, understand the task, by reading the intro and thinking. Write down:
- Who are the “users" of the algorithm?
    - Where is the result is going to be used?
    - What is important for the downstream task?
    - How one can measure it?
    - Where the data comes from in production?
    - Can we avoid solving this task at all?

Case study: image matching

Task understanding (1)



Noname manuscript No.
(will be inserted by the editor)

Image Matching Across Wide Baselines: From Paper to Practice

Yuhe Jin · Dmytro Mishkin · Anastasiia Mishchuk · Jiri Matas · Pascal Fua · Kwang Moo Yi · Eduard Trulls

arXiv:2003.01587v5 [cs.CV] 11 Feb 2021

Received: date / Accepted: date

Abstract We introduce a comprehensive benchmark for local features and robust estimation algorithms, focusing on the downstream task – the accuracy of the reconstructed camera pose – as our primary metric. Our pipeline's modular structure allows easy integration, configuration, and combination of different methods and heuristics. This is demonstrated by embedding dozens of popular algorithms and evaluating them, from seminal works to the cutting edge of machine learning research. We show that with proper settings, classical solutions may still outperform the *perceived state of the art*.

Besides establishing the *actual state of the art*, the conducted experiments reveal unexpected properties of Structure from Motion (SfM) pipelines that can help improve their performance, for both algorithmic and learned methods. Data and code are online[1], providing an easy-to-use and flexible framework for the benchmarking of local features and robust estimation methods, both *alongside* and *against* top-performing methods. This work provides a basis for the Image Matching Challenge[2].

Fig. 1 Every paper claims to outperform the state of the art. Is this possible, or an artifact of insufficient validation? On the left, we show stereo matches obtained with **D2-Net** (2019) [17], a state-of-the-art local feature, using OpenCV RANSAC with its default settings. We color the inliers in green if they are correct and in red otherwise. On the right, we show SIFT (1999) [54] with a carefully tuned MAGSAC [15] – notice how the latter performs much better. This illustrates our take-home message: to correctly evaluate a method's performance, it needs to be embedded within the pipeline used to solve a given problem, and the different components in said pipeline need to be tuned carefully and jointly, which requires engineering and domain expertise. We fill this need with a new, modular benchmark for sparse image matching, incorporating dozens of built-in methods.

This work was partially supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant "Deep Visual Geometry Machines" (RGPIN-2018-03788), by systems supplied by Compute Canada, and by Google's Visual Positioning Service. DM and JM were supported by OP VVV funded project CZ.02.1.01/0.0/0.0/16 019/0000765 "Research Center for Informatics". DM was also supported by CTU student grant SGS17/185/OHK3/3T/13 and by the Austrian Ministry for Transport, Innovation and Technology, the Federal Ministry for Digital and Economic Affairs, and the Province of Upper Austria in the frame of the COMET center SCCH. AM was supported by the Swiss National Science Foundation.

Y. Jin, K.M. Yi
Visual Computing Group, University of Victoria
E-mail: {yuhejin, kyi}@uvic.ca

A. Mishchuk, P. Fua
Computer Vision Lab, École Polytechnique Fédérale de Lausanne
E-mail: {anastasiia.mishchuk, pascal.fua}@epfl.ch

D. Mishkin, J. Matas
Center for Machine Perception, Czech Technical University in Prague
E-mail: {mishkdmy, matas}@cmp.felk.cvut.cz

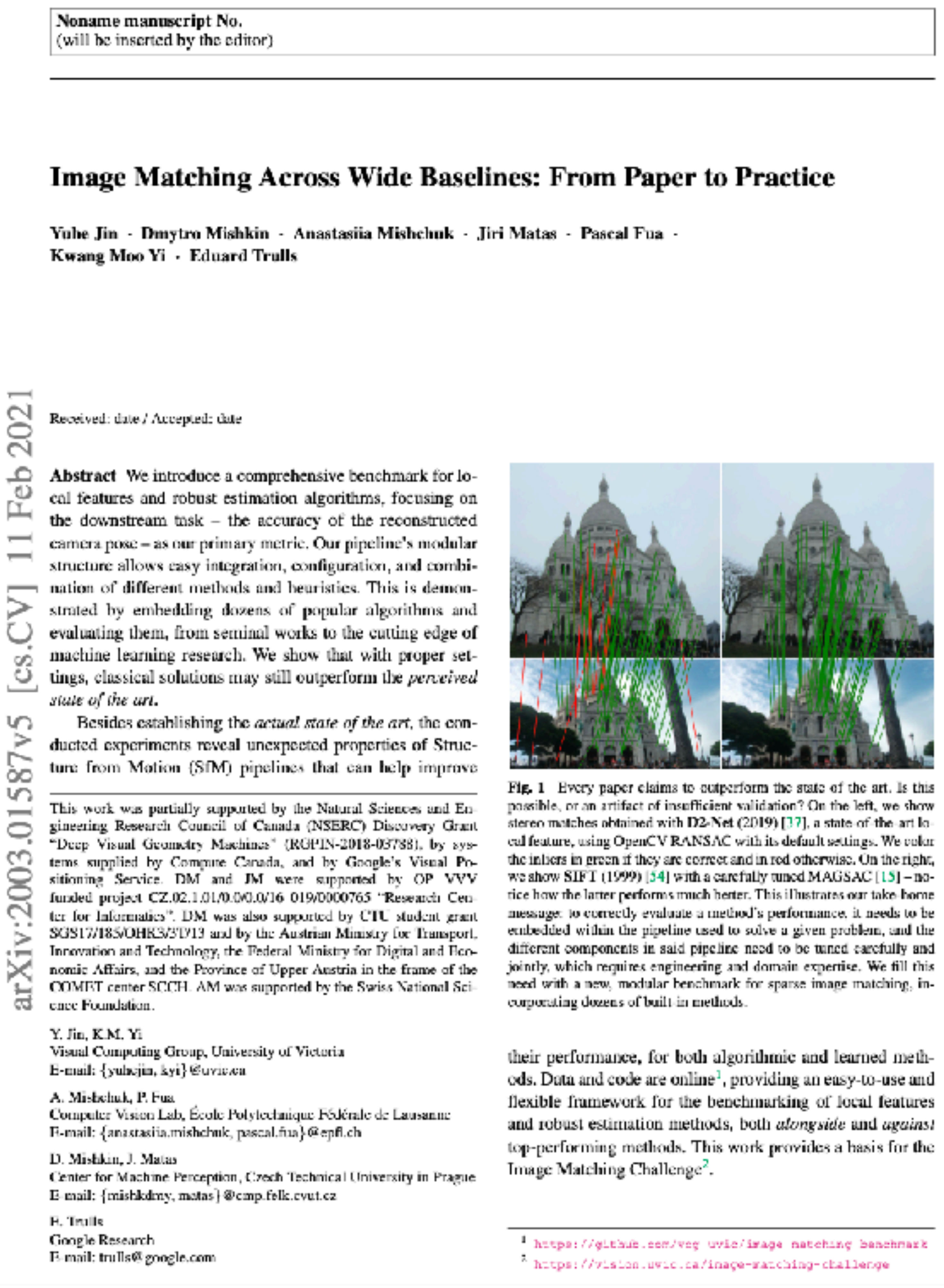
E. Trulls
Google Research
E-mail: trulls@google.com

[1] https://github.com/vcg-uvic/image-matching-benchmark
[2] https://vision.uvic.ca/image-matching-challenge

Case study: image matching

Task understanding (1)

- What is the task?



Noname manuscript No.
(will be inserted by the editor)

Image Matching Across Wide Baselines: From Paper to Practice

Yuhe Jin · Dmytro Mishkin · Anastasiia Mishchuk · Jiri Matas · Pascal Fua · Kwang Moo Yi · Eduard Trulls

Received: date / Accepted: date

Abstract We introduce a comprehensive benchmark for local features and robust estimation algorithms, focusing on the downstream task – the accuracy of the reconstructed camera pose – as our primary metric. Our pipeline's modular structure allows easy integration, configuration, and combination of different methods and heuristics. This is demonstrated by embedding dozens of popular algorithms and evaluating them, from seminal works to the cutting edge of machine learning research. We show that with proper settings, classical solutions may still outperform the *perceived state of the art*.

Besides establishing the *actual state of the art*, the conducted experiments reveal unexpected properties of Structure from Motion (SfM) pipelines that can help improve their performance, for both algorithmic and learned methods. Data and code are online[1], providing an easy-to-use and flexible framework for the benchmarking of local features and robust estimation methods, both *alongside* and *against* top-performing methods. This work provides a basis for the Image Matching Challenge[2].

Fig. 1 Every paper claims to outperform the state of the art. Is this possible, or an artifact of insufficient validation? On the left, we show stereo matches obtained with D2-Net (2019) [17], a state-of-the-art local feature, using OpenCV RANSAC with its default settings. We color the inliers in green if they are correct and in red otherwise. On the right, we show SIFT (1999) [54] with a carefully tuned MAGSAC [15] – notice how the latter performs much better. This illustrates our take-home message: to correctly evaluate a method's performance, it needs to be embedded within the pipeline used to solve a given problem, and the different components in said pipeline need to be tuned carefully and jointly, which requires engineering and domain expertise. We fill this need with a new, modular benchmark for sparse image matching, incorporating dozens of built-in methods.

This work was partially supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant "Deep Visual Geometry Machines" (RGPIN-2018-03788), by systems supplied by Compute Canada, and by Google's Visual Positioning Service. DM and JM were supported by OP VVV funded project CZ.02.1.01/0.0/0.0/16 019/0000765 "Research Center for Informatics". DM was also supported by CTU student grant SGS17/185/OHK3/3T/13 and by the Austrian Ministry for Transport, Innovation and Technology, the Federal Ministry for Digital and Economic Affairs, and the Province of Upper Austria in the frame of the COMET center SCCH. AM was supported by the Swiss National Science Foundation.

Y. Jin, K.M. Yi
Visual Computing Group, University of Victoria
E-mail: {yuhejin, kyi}@uvic.ca

A. Mishchuk, P. Fua
Computer Vision Lab, École Polytechnique Fédérale de Lausanne
E-mail: {anastasiia.mishchuk, pascal.fua}@epfl.ch

D. Mishkin, J. Matas
Center for Machine Perception, Czech Technical University in Prague
E-mail: {mishkdmy, matas}@cmp.felk.cvut.cz

E. Trulls
Google Research
E-mail: trulls@google.com

[1] https://github.com/vcg-uvic/image-matching-benchmark
[2] https://vision.uvic.ca/image-matching-challenge

Case study: image matching

Task understanding (1)

- What is the task?
  - Image matching.

Noname manuscript No.
(will be inserted by the editor)

Image Matching Across Wide Baselines: From Paper to Practice

Yuhe Jin · Dmytro Mishkin · Anastasiia Mishchuk · Jiri Matas · Pascal Fua · Kwang Moo Yi · Eduard Trulls

arXiv:2003.01587v5 [cs.CV] 11 Feb 2021

Received: date / Accepted: date

Abstract We introduce a comprehensive benchmark for local features and robust estimation algorithms, focusing on the downstream task – the accuracy of the reconstructed camera pose – as our primary metric. Our pipeline's modular structure allows easy integration, configuration, and combination of different methods and heuristics. This is demonstrated by embedding dozens of popular algorithms and evaluating them, from seminal works to the cutting edge of machine learning research. We show that with proper settings, classical solutions may still outperform the *perceived state of the art*.

Besides establishing the *actual state of the art*, the conducted experiments reveal unexpected properties of Structure from Motion (SfM) pipelines that can help improve their performance, for both algorithmic and learned methods. Data and code are online[1], providing an easy-to-use and flexible framework for the benchmarking of local features and robust estimation methods, both *alongside* and *against* top-performing methods. This work provides a basis for the Image Matching Challenge[2].

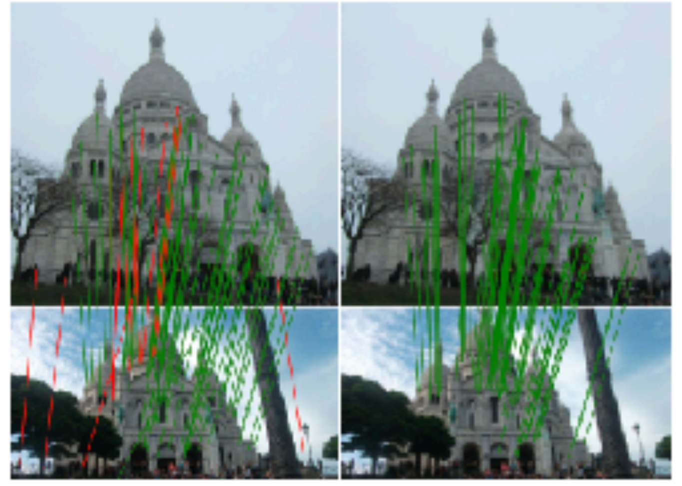


Fig. 1 Every paper claims to outperform the state of the art. Is this possible, or an artifact of insufficient validation? On the left, we show stereo matches obtained with **D2-Net** (2019) [17], a state of the art local feature, using OpenCV RANSAC with its default settings. We color the inliers in green if they are correct and in red otherwise. On the right, we show SIFT (1999) [54] with a carefully tuned MAGSAC [15] – notice how the latter performs much better. This illustrates our take-home message: to correctly evaluate a method's performance, it needs to be embedded within the pipeline used to solve a given problem, and the different components in said pipeline need to be tuned carefully and jointly, which requires engineering and domain expertise. We fill this need with a new, modular benchmark for sparse image matching, incorporating dozens of built-in methods.

This work was partially supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant "Deep Visual Geometry Machines" (RGPIN-2018-03788), by systems supplied by Compute Canada, and by Google's Visual Positioning Service. DM and JM were supported by OP VVV funded project CZ.02.1.01/0.0/0.0/16 019/0000765 "Research Center for Informatics". DM was also supported by CTU student grant SGS17/185/OHK3/3T/13 and by the Austrian Ministry for Transport, Innovation and Technology, the Federal Ministry for Digital and Economic Affairs, and the Province of Upper Austria in the frame of the COMET center SCCH. AM was supported by the Swiss National Science Foundation.

Y. Jin, K.M. Yi
Visual Computing Group, University of Victoria
E-mail: {yuhejin, kyi}@uvic.ca

A. Mishchuk, P. Fua
Computer Vision Lab, École Polytechnique Fédérale de Lausanne
E-mail: {anastasiia.mishchuk, pascal.fua}@epfl.ch

D. Mishkin, J. Matas
Center for Machine Perception, Czech Technical University in Prague
E-mail: {mishkdmy, matas}@cmp.felk.cvut.cz

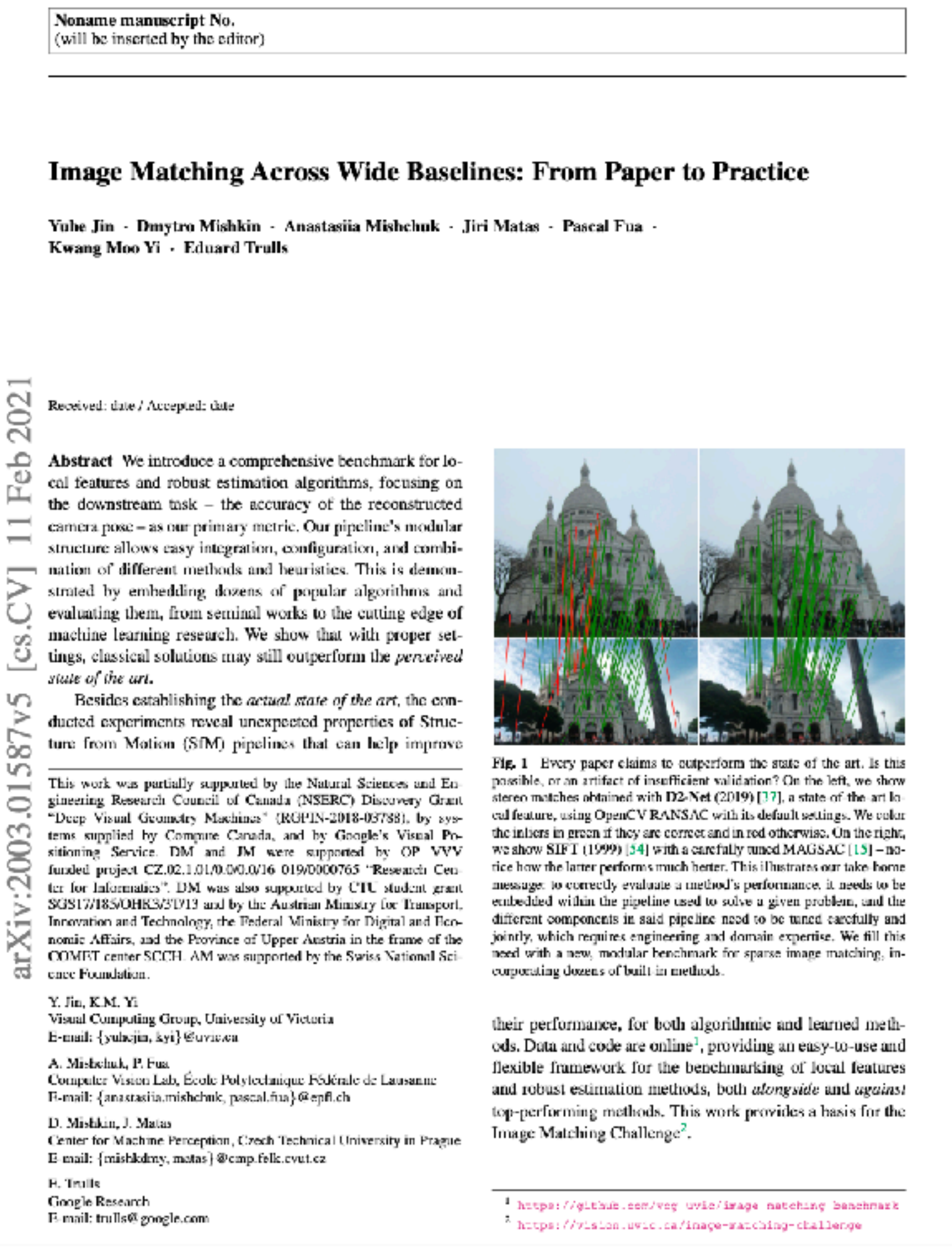
E. Trulls
Google Research
E-mail: trulls@google.com

[1] https://github.com/vcg-uvic/image-matching-benchmark
[2] https://vision.uvic.ca/image-matching-challenge

Case study: image matching

Task understanding (1)

- What is the task?
    - Image matching.
- What are the inputs and output of the task?



Noname manuscript No.
(will be inserted by the editor)

Image Matching Across Wide Baselines: From Paper to Practice

Yuhe Jin · Dmytro Mishkin · Anastasiia Mishchuk · Jiri Matas · Pascal Fua · Kwang Moo Yi · Eduard Trulls

arXiv:2003.01587v5 [cs.CV] 11 Feb 2021

Received: date / Accepted: date

Abstract We introduce a comprehensive benchmark for local features and robust estimation algorithms, focusing on the downstream task – the accuracy of the reconstructed camera pose – as our primary metric. Our pipeline's modular structure allows easy integration, configuration, and combination of different methods and heuristics. This is demonstrated by embedding dozens of popular algorithms and evaluating them, from seminal works to the cutting edge of machine learning research. We show that with proper settings, classical solutions may still outperform the *perceived state of the art*.

Besides establishing the *actual state of the art*, the conducted experiments reveal unexpected properties of Structure from Motion (SfM) pipelines that can help improve their performance, for both algorithmic and learned methods. Data and code are online[1], providing an easy-to-use and flexible framework for the benchmarking of local features and robust estimation methods, both *alongside* and *against* top-performing methods. This work provides a basis for the Image Matching Challenge[2].

Fig. 1 Every paper claims to outperform the state of the art. Is this possible, or an artifact of insufficient validation? On the left, we show stereo matches obtained with D2-Net (2019) [17], a state-of-the-art local feature, using OpenCV RANSAC with its default settings. We color the inliers in green if they are correct and in red otherwise. On the right, we show SIFT (1999) [54] with a carefully tuned MAGSAC [15] – notice how the latter performs much better. This illustrates our take-home message: to correctly evaluate a method's performance, it needs to be embedded within the pipeline used to solve a given problem, and the different components in said pipeline need to be tuned carefully and jointly, which requires engineering and domain expertise. We fill this need with a new, modular benchmark for sparse image matching, incorporating dozens of built-in methods.

This work was partially supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant "Deep Visual Geometry Machines" (RGPIN-2018-03788), by systems supplied by Compute Canada, and by Google's Visual Positioning Service. DM and JM were supported by OP VVV funded project CZ.02.1.01/0.0/0.0/16 019/0000765 "Research Center for Informatics". DM was also supported by CTU student grant SGS17/185/OHK3/3T/13 and by the Austrian Ministry for Transport, Innovation and Technology, the Federal Ministry for Digital and Economic Affairs, and the Province of Upper Austria in the frame of the COMET center SCCH. AM was supported by the Swiss National Science Foundation.

Y. Jin, K.M. Yi
Visual Computing Group, University of Victoria
E-mail: {yuhejin, kyi}@uvic.ca

A. Mishchuk, P. Fua
Computer Vision Lab, École Polytechnique Fédérale de Lausanne
E-mail: {anastasiia.mishchuk, pascal.fua}@epfl.ch

D. Mishkin, J. Matas
Center for Machine Perception, Czech Technical University in Prague
E-mail: {mishkdmy, matas}@cmp.felk.cvut.cz

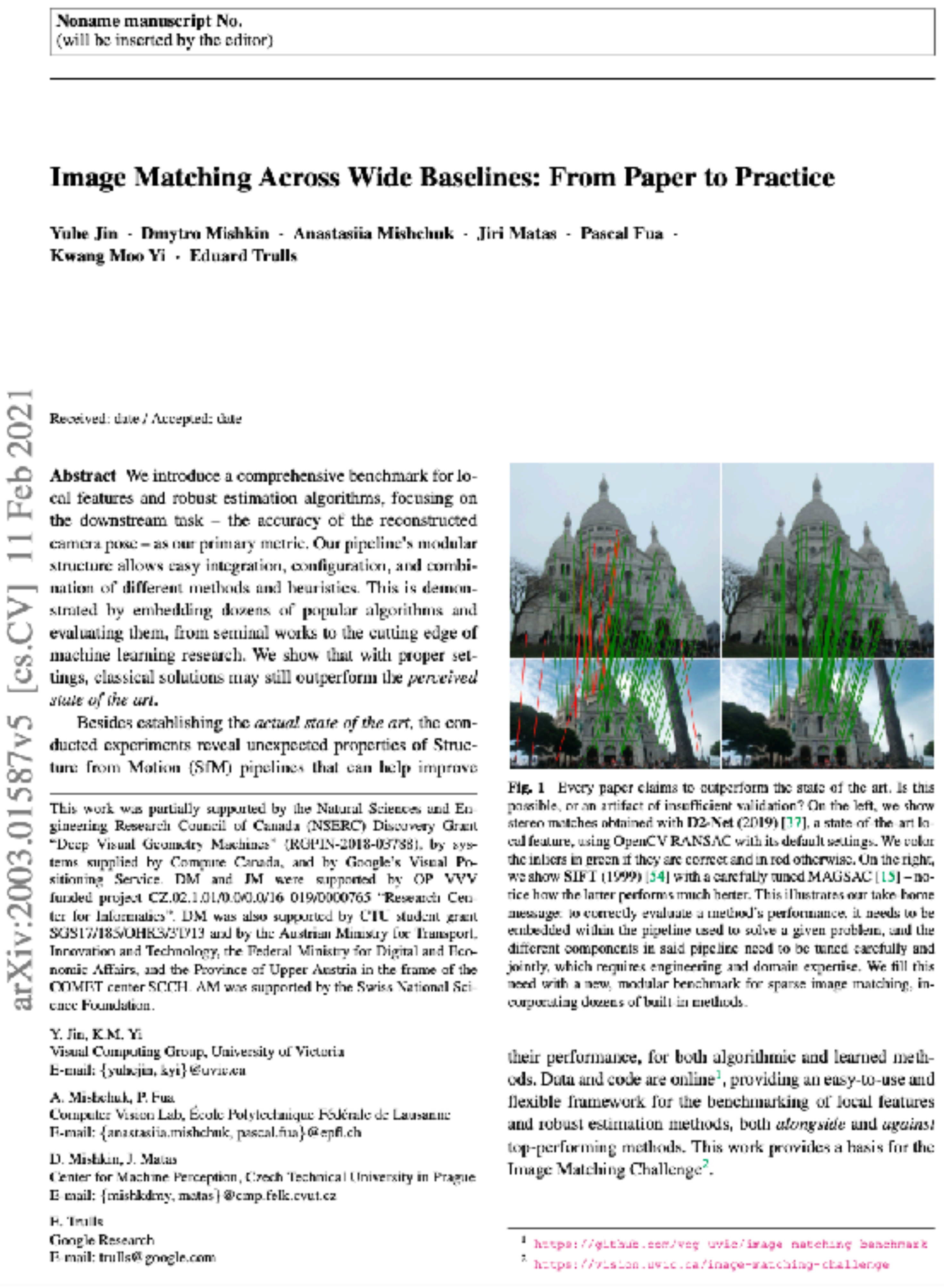
E. Trulls
Google Research
E-mail: trulls@google.com

[1] https://github.com/vcg-uvic/image-matching-benchmark
[2] https://vision.uvic.ca/image-matching-challenge

Case study: image matching

Task understanding (1)

- What is the task?
    - Image matching.
- What are the inputs and output of the task?
    - Missing in the intro! Need to google



Noname manuscript No.
(will be inserted by the editor)

Image Matching Across Wide Baselines: From Paper to Practice

Yuhe Jin · Dmytro Mishkin · Anastasiia Mishchuk · Jiri Matas · Pascal Fua · Kwang Moo Yi · Eduard Trulls

arXiv:2003.01587v5 [cs.CV] 11 Feb 2021

Received: date / Accepted: date

Abstract We introduce a comprehensive benchmark for local features and robust estimation algorithms, focusing on the downstream task – the accuracy of the reconstructed camera pose – as our primary metric. Our pipeline's modular structure allows easy integration, configuration, and combination of different methods and heuristics. This is demonstrated by embedding dozens of popular algorithms and evaluating them, from seminal works to the cutting edge of machine learning research. We show that with proper settings, classical solutions may still outperform the *perceived state of the art.*

Besides establishing the *actual state of the art*, the conducted experiments reveal unexpected properties of Structure from Motion (SfM) pipelines that can help improve their performance, for both algorithmic and learned methods. Data and code are online[1], providing an easy-to-use and flexible framework for the benchmarking of local features and robust estimation methods, both *alongside* and *against* top-performing methods. This work provides a basis for the Image Matching Challenge[2].

Fig. 1 Every paper claims to outperform the state of the art. Is this possible, or an artifact of insufficient validation? On the left, we show stereo matches obtained with D2-Net (2019) [17], a state-of-the-art local feature, using OpenCV RANSAC with its default settings. We color the inliers in green if they are correct and in red otherwise. On the right, we show SIFT (1999) [54] with a carefully tuned MAGSAC [15] – notice how the latter performs much better. This illustrates our take-home message: to correctly evaluate a method's performance, it needs to be embedded within the pipeline used to solve a given problem, and the different components in said pipeline need to be tuned carefully and jointly, which requires engineering and domain expertise. We fill this need with a new, modular benchmark for sparse image matching, incorporating dozens of built-in methods.

This work was partially supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant "Deep Visual Geometry Machines" (RGPIN-2018-03788), by systems supplied by Compute Canada, and by Google's Visual Positioning Service. DM and JM were supported by OP VVV funded project CZ.02.1.01/0.0/0.0/16 019/0000765 "Research Center for Informatics". DM was also supported by CTU student grant SGS17/185/OHK3/3T/13 and by the Austrian Ministry for Transport, Innovation and Technology, the Federal Ministry for Digital and Economic Affairs, and the Province of Upper Austria in the frame of the COMET center SCCH. AM was supported by the Swiss National Science Foundation.

Y. Jin, K.M. Yi
Visual Computing Group, University of Victoria
E-mail: {yuhejin, kyi}@uvic.ca

A. Mishchuk, P. Fua
Computer Vision Lab, École Polytechnique Fédérale de Lausanne
E-mail: {anastasiia.mishchuk, pascal.fua}@epfl.ch

D. Mishkin, J. Matas
Center for Machine Perception, Czech Technical University in Prague
E-mail: {mishkdmy, matas}@cmp.felk.cvut.cz

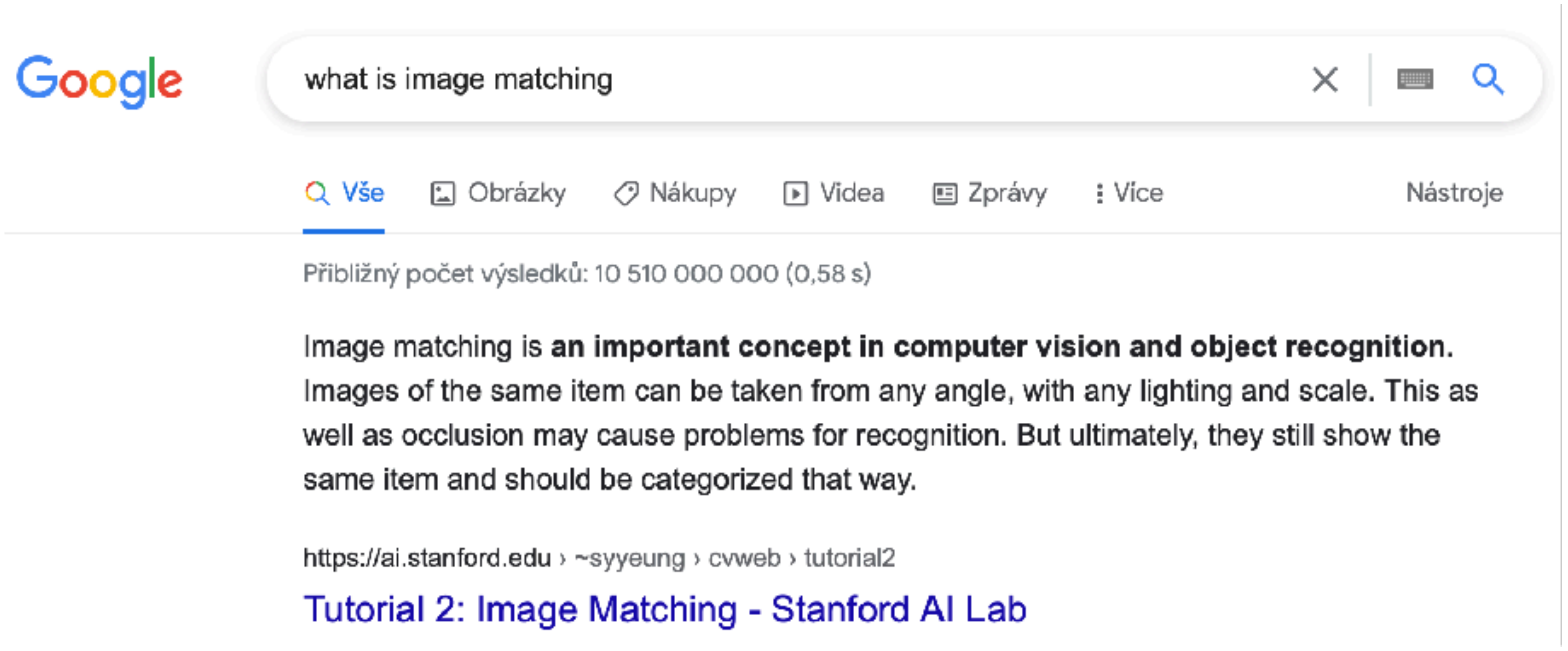
E. Trulls
Google Research
E-mail: trulls@google.com

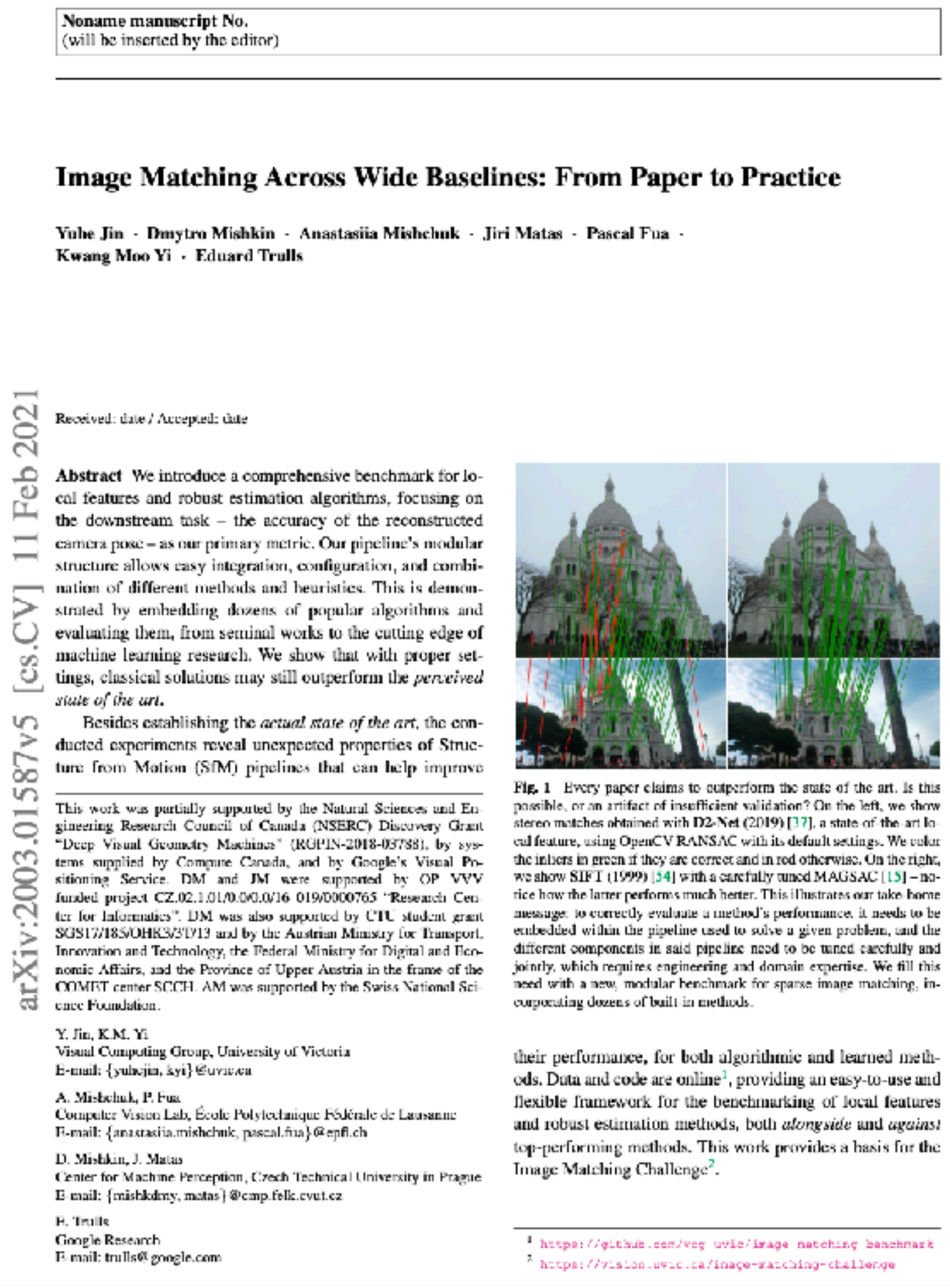
1 https://github.com/vcg-uvic/image-matching-benchmark
2 https://vision.uvic.ca/image-matching-challenge

Case study: image matching

Task understanding (1)

- What is the task?
    - Image matching.
- What are the inputs and output of the task?
    - Missing in the intro! Need to google





Noname manuscript No.
(will be inserted by the editor)

Image Matching Across Wide Baselines: From Paper to Practice

Yuhe Jin · Dmytro Mishkin · Anastasiia Mishchuk · Jiri Matas · Pascal Fua · Kwang Moo Yi · Eduard Trulls

arXiv:2003.01587v5 [cs.CV] 11 Feb 2021

Received: date / Accepted: date

Abstract We introduce a comprehensive benchmark for local features and robust estimation algorithms, focusing on the downstream task – the accuracy of the reconstructed camera pose – as our primary metric. Our pipeline's modular structure allows easy integration, configuration, and combination of different methods and heuristics. This is demonstrated by embedding dozens of popular algorithms and evaluating them, from seminal works to the cutting edge of machine learning research. We show that with proper settings, classical solutions may still outperform the *perceived state of the art*.

Besides establishing the *actual state of the art*, the conducted experiments reveal unexpected properties of Structure from Motion (SfM) pipelines that can help improve their performance, for both algorithmic and learned methods. Data and code are online[1], providing an easy-to-use and flexible framework for the benchmarking of local features and robust estimation methods, both *alongside* and *against* top-performing methods. This work provides a basis for the Image Matching Challenge[2].

Fig. 1 Every paper claims to outperform the state of the art. Is this possible, or an artifact of insufficient validation? On the left, we show stereo matches obtained with D2-Net (2019) [37], a state-of-the-art local feature, using OpenCV RANSAC with its default settings. We color the inliers in green if they are correct and in red otherwise. On the right, we show SIFT (1999) [54] with a carefully tuned MAGSAC [15] – notice how the latter performs much better. This illustrates our take-home message: to correctly evaluate a method's performance, it needs to be embedded within the pipeline used to solve a given problem, and the different components in said pipeline need to be tuned carefully and jointly, which requires engineering and domain expertise. We fill this need with a new, modular benchmark for sparse image matching, incorporating dozens of built-in methods.

This work was partially supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant "Deep Visual Geometry Machines" (RGPIN-2018-03788), by systems supplied by Compute Canada, and by Google's Visual Positioning Service. DM and JM were supported by OP VVV funded project CZ.02.1.01/0.0/0.0/16 019/0000765 "Research Center for Informatics". DM was also supported by CTU student grant SGS17/185/OHK3/3T/13 and by the Austrian Ministry for Transport, Innovation and Technology, the Federal Ministry for Digital and Economic Affairs, and the Province of Upper Austria in the frame of the COMET center SCCH. AM was supported by the Swiss National Science Foundation.

Y. Jin, K.M. Yi
Visual Computing Group, University of Victoria
E-mail: {yuhejin, kyi}@uvic.ca

A. Mishchuk, P. Fua
Computer Vision Lab, École Polytechnique Fédérale de Lausanne
E-mail: {anastasiia.mishchuk, pascal.fua}@epfl.ch

D. Mishkin, J. Matas
Center for Machine Perception, Czech Technical University in Prague
E-mail: {mishkdmy, matas}@cmp.felk.cvut.cz

E. Trulls
Google Research
E-mail: trulls@google.com

[1] https://github.com/vcg-uvic/image-matching-benchmark
[2] https://vision.uvic.ca/image-matching-challenge

Case study: image matching

Task understanding (1)

- Inputs: images of the “same item”.
- Output (after more googling): correspondences and relative camera pose
- Note: people can use different terms for the same thing:
  - "Image matching” == “Wide baseline stereo"

The wide multiple baseline stereo (WxBS) is **a process of establishing a sufficient number of pixel or region correspondences from two or more images depicting the same scene to estimate the geometric relationship between cameras, which produced these images**. Mar 27, 2020

https://ducha-aiki.github.io › 2020/03/27 › intro

The Role of Wide Baseline Stereo in the Deep Learning World

Case study: image matching

Task understanding (2)

Case study: image matching

Task understanding (2)

- Who are the “users" of the algorithm? Where is the result is going to be used?

Case study: image matching

Task understanding (2)

- Who are the “users" of the algorithm? Where is the result is going to be used?

1 Introduction

Matching two or more views of a scene is at the core of fundamental computer vision problems, including image retrieval [54,7,78,104,70], 3D reconstruction [3,47,90,122], re-localization [85,86,57], and SLAM [68,33,34]. Despite

Case study: image matching

Task understanding (2)

- Who are the “users" of the algorithm? Where is the result is going to be used?
  - Image retrieval

1 Introduction

Matching two or more views of a scene is at the core of fundamental computer vision problems, including image retrieval [54,7,78,104,70], 3D reconstruction [3,47,90,122], re-localization [85,86,57], and SLAM [68,33,34]. Despite

Case study: image matching

Task understanding (2)

- Who are the “users" of the algorithm? Where is the result is going to be used?
    - Image retrieval
    - 3D reconstruction

1 Introduction

Matching two or more views of a scene is at the core of fundamental computer vision problems, including image retrieval [54,7,78,104,70], 3D reconstruction [3,47,90,122], re-localization [85,86,57], and SLAM [68,33,34]. Despite

Case study: image matching
Task understanding (2)

- Who are the “users" of the algorithm? Where is the result is going to be used?
  - Image retrieval
  - 3D reconstruction
  - visual localization

1 Introduction

Matching two or more views of a scene is at the core of fundamental computer vision problems, including image retrieval [54,7,78,104,70], 3D reconstruction [3,47,90,122], re-localization [85,86,57], and SLAM [68,33,34]. Despite

Case study: image matching

Task understanding (2)

- Who are the “users" of the algorithm? Where is the result is going to be used?
  - Image retrieval
  - 3D reconstruction
  - visual localization
  - SLAM

1 Introduction

Matching two or more views of a scene is at the core of fundamental computer vision problems, including image retrieval [54,7,78,104,70], 3D reconstruction [3,47,90,122], re-localization [85,86,57], and SLAM [68,33,34]. Despite

Case study: image matching

Task understanding (2)

- Who are the “users" of the algorithm? Where is the result is going to be used?
  - Image retrieval
  - 3D reconstruction
  - visual localization
  - SLAM

1 Introduction

Matching two or more views of a scene is at the core of fundamental computer vision problems, including image retrieval [54,7,78,104,70], 3D reconstruction [3,47,90,122], re-localization [85,86,57], and SLAM [68,33,34]. Despite

- Now we may, or may not need to check about this “users"

Case study: image matching

Task understanding (2)

Case study: image matching

Task understanding (2)

- What is important for the downstream task?

Case study: image matching
Task understanding (2)

- What is important for the downstream task?
  - Image retrieval — if the result is relevant to query

Case study: image matching

Task understanding (2)

- What is important for the downstream task?
  - Image retrieval — if the result is relevant to query
  - 3D reconstruction — accuracy and completeness of the reconstruction

Case study: image matching
Task understanding (2)

- What is important for the downstream task?
  - Image retrieval — if the result is relevant to query
  - 3D reconstruction — accuracy and completeness of the reconstruction
  - visual localization — camera pose accuracy, robustness to nuisance

Case study: image matching
Task understanding (2)

- What is important for the downstream task?
  - Image retrieval — if the result is relevant to query
  - 3D reconstruction — accuracy and completeness of the reconstruction
  - visual localization — camera pose accuracy, robustness to nuisance
  - SLAM — camera pose accuracy, realtime speed, blur-robustness

Case study: image matching

Task understanding (2)

- What is important for the downstream task?
  - Image retrieval — if the result is relevant to query
  - 3D reconstruction — accuracy and completeness of the reconstruction
  - visual localization — camera pose accuracy, robustness to nuisance
  - SLAM — camera pose accuracy, realtime speed, blur-robustness

Is it relevant to me?

Case study: image matching

Task understanding (3)

Case study: image matching

Task understanding (3)

- We have already googled and checked other papers to answer the questions, which are necessary to understand the paper we are reading now

Case study: image matching
Task understanding (3)

- We have already googled and checked other papers to answer the questions, which are necessary to understand the paper we are reading now
- It is OK to deviate (a lot), when you are reading

Case study: image matching
Task understanding (3)

- We have already googled and checked other papers to answer the questions, which are necessary to understand the paper we are reading now
- It is OK to deviate (a lot), when you are reading
- Especially when you are new to the area

Case study: image matching

Task understanding (3)

- We have already googled and checked other papers to answer the questions, which are necessary to understand the paper we are reading now
- It is OK to deviate (a lot), when you are reading
- Especially when you are new to the area
- It is a good idea to “learn in public” and write answers to the question in a blog post.

Case study: image matching
How can we measure that?

- What is important for the downstream task?
  - Image retrieval — if the result is relevant to query
  - 3D reconstruction — accuracy and completeness of the reconstruction
  - visual localization — camera pose accuracy, robustness to nuisance
  - SLAM — camera pose accuracy, realtime speed, blur-robustness

Case study: image matching

How can we measure that?

- What is important for the downstream task?
  - Image retrieval — if the result is relevant to query
  - 3D reconstruction — accuracy and completeness of the reconstruction
  - visual localization — camera pose accuracy, robustness to nuisance
  - SLAM — camera pose accuracy, realtime speed, blur-robustness

Recall and precision

Case study: image matching

How can we measure that?

- What is important for the downstream task?
  - Image retrieval — if the result is relevant to query
  - 3D reconstruction — accuracy and completeness of the reconstruction
  - visual localization — camera pose accuracy, robustness to nuisance
  - SLAM — camera pose accuracy, realtime speed, blur-robustness

Recall and precision

Lots of metrics

Case study: image matching

How can we measure that?

- What is important for the downstream task?
  - Image retrieval — if the result is relevant to query
  - 3D reconstruction — accuracy and completeness of the reconstruction
  - visual localization — camera pose accuracy, robustness to nuisance
  - SLAM — camera pose accuracy, realtime speed, blur-robustness

Recall and precision

Lots of metrics

Pose error, accuracy

Case study: image matching

How can we measure that?

- What is important for the downstream task?
  - Image retrieval — if the result is relevant to query — Recall and precision
  - 3D reconstruction — accuracy and completeness of the reconstruction — Lots of metrics
  - visual localization — camera pose accuracy, robustness to nuisance — Pose error, accuracy
  - SLAM — camera pose accuracy, realtime speed, blur-robustness — Pose error, fps

Case study: image matching
Task understanding (4)

Case study: image matching
Task understanding (4)

- Where the data comes from in production?

Case study: image matching

Task understanding (4)

- Where the data comes from in production?
  - Image retrieval — user upload, can be anything

Case study: image matching

Task understanding (4)

- Where the data comes from in production?
  - Image retrieval — user upload, can be anything
  - 3D reconstruction — professional cameras, drones, smartphone, historical photos, etc

Case study: image matching
Task understanding (4)

- Where the data comes from in production?
  - Image retrieval — user upload, can be anything
  - 3D reconstruction — professional cameras, drones, smartphone, historical photos, etc
  - visual localization — robot camera, car, smartphone, internet photo (OSI)

Case study: image matching
Task understanding (4)

- Where the data comes from in production?
  - Image retrieval — user upload, can be anything
  - 3D reconstruction — professional cameras, drones, smartphone, historical photos, etc
  - visual localization — robot camera, car, smartphone, internet photo (OSI)
  - SLAM — robot, UAV, smart glasses, car

Back to the paper

What does the paper evaluate?

Case study: Image matching

- What does paper measure?
- Where the data in the paper comes from?
- Where the ground truth labels come from?

What does the paper evaluate?

Again, we are speaking purely about the task now

3 The Image Matching Challenge PhotoTourism Dataset

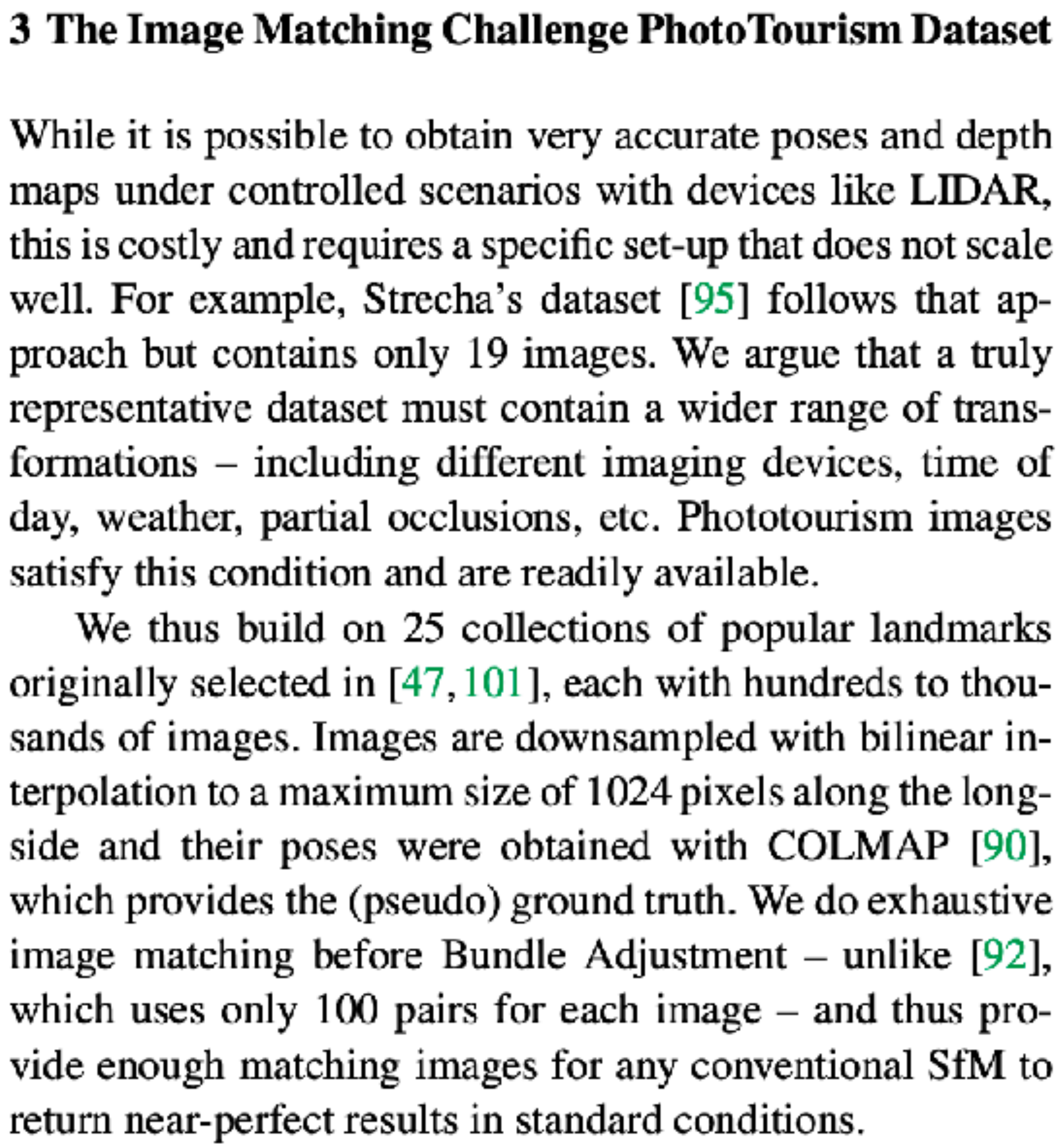
While it is possible to obtain very accurate poses and depth maps under controlled scenarios with devices like LIDAR, this is costly and requires a specific set-up that does not scale well. For example, Strecha's dataset [95] follows that approach but contains only 19 images. We argue that a truly representative dataset must contain a wider range of transformations – including different imaging devices, time of day, weather, partial occlusions, etc. Phototourism images satisfy this condition and are readily available.

We thus build on 25 collections of popular landmarks originally selected in [47, 101], each with hundreds to thousands of images. Images are downsampled with bilinear interpolation to a maximum size of 1024 pixels along the long-side and their poses were obtained with COLMAP [90], which provides the (pseudo) ground truth. We do exhaustive image matching before Bundle Adjustment – unlike [92], which uses only 100 pairs for each image – and thus provide enough matching images for any conventional SfM to return near-perfect results in standard conditions.

What does the paper evaluate?

Again, we are speaking purely about the task now

- What does paper measure — mean average accuracy



3 The Image Matching Challenge PhotoTourism Dataset

While it is possible to obtain very accurate poses and depth maps under controlled scenarios with devices like LIDAR, this is costly and requires a specific set-up that does not scale well. For example, Strecha's dataset [95] follows that approach but contains only 19 images. We argue that a truly representative dataset must contain a wider range of transformations – including different imaging devices, time of day, weather, partial occlusions, etc. Phototourism images satisfy this condition and are readily available.

We thus build on 25 collections of popular landmarks originally selected in [47, 101], each with hundreds to thousands of images. Images are downsampled with bilinear interpolation to a maximum size of 1024 pixels along the long-side and their poses were obtained with COLMAP [90], which provides the (pseudo) ground truth. We do exhaustive image matching before Bundle Adjustment – unlike [92], which uses only 100 pairs for each image – and thus provide enough matching images for any conventional SfM to return near-perfect results in standard conditions.

What does the paper evaluate?

Again, we are speaking purely about the task now

- What does paper measure — mean average accuracy
- Where the data in the paper comes from?

3 The Image Matching Challenge PhotoTourism Dataset

While it is possible to obtain very accurate poses and depth maps under controlled scenarios with devices like LIDAR, this is costly and requires a specific set-up that does not scale well. For example, Strecha's dataset [95] follows that approach but contains only 19 images. We argue that a truly representative dataset must contain a wider range of transformations – including different imaging devices, time of day, weather, partial occlusions, etc. Phototourism images satisfy this condition and are readily available.

We thus build on 25 collections of popular landmarks originally selected in [47, 101], each with hundreds to thousands of images. Images are downsampled with bilinear interpolation to a maximum size of 1024 pixels along the long-side and their poses were obtained with COLMAP [90], which provides the (pseudo) ground truth. We do exhaustive image matching before Bundle Adjustment – unlike [92], which uses only 100 pairs for each image – and thus provide enough matching images for any conventional SfM to return near-perfect results in standard conditions.

What does the paper evaluate?

Again, we are speaking purely about the task now

- What does paper measure — mean average accuracy
- Where the data in the paper comes from?
  - YCC100M dataset, from flickr (user photos)

3 The Image Matching Challenge PhotoTourism Dataset

While it is possible to obtain very accurate poses and depth maps under controlled scenarios with devices like LIDAR, this is costly and requires a specific set-up that does not scale well. For example, Strecha's dataset [95] follows that approach but contains only 19 images. We argue that a truly representative dataset must contain a wider range of transformations – including different imaging devices, time of day, weather, partial occlusions, etc. Phototourism images satisfy this condition and are readily available.

We thus build on 25 collections of popular landmarks originally selected in [47, 101], each with hundreds to thousands of images. Images are downsampled with bilinear interpolation to a maximum size of 1024 pixels along the long-side and their poses were obtained with COLMAP [90], which provides the (pseudo) ground truth. We do exhaustive image matching before Bundle Adjustment – unlike [92], which uses only 100 pairs for each image – and thus provide enough matching images for any conventional SfM to return near-perfect results in standard conditions.

What does the paper evaluate?

Again, we are speaking purely about the task now

- What does paper measure — mean average accuracy
- Where the data in the paper comes from?
    - YCC100M dataset, from flickr (user photos)
- Where the ground truth labels come from?

3 The Image Matching Challenge PhotoTourism Dataset

While it is possible to obtain very accurate poses and depth maps under controlled scenarios with devices like LIDAR, this is costly and requires a specific set-up that does not scale well. For example, Strecha's dataset [95] follows that approach but contains only 19 images. We argue that a truly representative dataset must contain a wider range of transformations – including different imaging devices, time of day, weather, partial occlusions, etc. Phototourism images satisfy this condition and are readily available.

We thus build on 25 collections of popular landmarks originally selected in [47, 101], each with hundreds to thousands of images. Images are downsampled with bilinear interpolation to a maximum size of 1024 pixels along the long-side and their poses were obtained with COLMAP [90], which provides the (pseudo) ground truth. We do exhaustive image matching before Bundle Adjustment – unlike [92], which uses only 100 pairs for each image – and thus provide enough matching images for any conventional SfM to return near-perfect results in standard conditions.

What does the paper evaluate?

Again, we are speaking purely about the task now

- What does paper measure — mean average accuracy
- Where the data in the paper comes from?
    - YCC100M dataset, from flickr (user photos)
- Where the ground truth labels come from?
    - COLMAP (SIFT-based 3d reconstruction software)

3 The Image Matching Challenge PhotoTourism Dataset

While it is possible to obtain very accurate poses and depth maps under controlled scenarios with devices like LIDAR, this is costly and requires a specific set-up that does not scale well. For example, Strecha's dataset [95] follows that approach but contains only 19 images. We argue that a truly representative dataset must contain a wider range of transformations – including different imaging devices, time of day, weather, partial occlusions, etc. Phototourism images satisfy this condition and are readily available.

We thus build on 25 collections of popular landmarks originally selected in [47, 101], each with hundreds to thousands of images. Images are downsampled with bilinear interpolation to a maximum size of 1024 pixels along the long-side and their poses were obtained with COLMAP [90], which provides the (pseudo) ground truth. We do exhaustive image matching before Bundle Adjustment – unlike [92], which uses only 100 pairs for each image – and thus provide enough matching images for any conventional SfM to return near-perfect results in standard conditions.

How the paper data & metric matches user?

Case study: image matching, metric

- User metrics:
  - Image retrieval — Recall and precision
  - 3D reconstruction — many metrics
  - visual localization — camera pose accuracy
  - SLAM — camera pose accuracy, realtime speed

How the paper data & metric matches user?

Case study: image matching, metric

- User metrics:
  - Image retrieval — Recall and precision
  - 3D reconstruction — many metrics
  - visual localization — camera pose accuracy
  - SLAM — camera pose accuracy, realtime speed

- Paper metrics:
  - mean average accuracy on camera pose

How the paper data & metric matches user?

Case study: image matching, metric

- User metrics:
  - Image retrieval — Recall and precision
  - 3D reconstruction — many metrics
  - visual localization — camera pose accuracy
  - SLAM — camera pose accuracy, realtime speed

✗

- Paper metrics:
  - mean average accuracy on camera pose

How the paper data & metric matches user?

Case study: image matching, metric

Note: not clear, why paper says that image matching is used in image retrieval
Add to the ToDo list to figure out (low priority)

- User metrics:
  - ~~Image retrieval — Recall and precision~~
  - 3D reconstruction — many metrics
  - visual localization — camera pose accuracy
  - SLAM — camera pose accuracy, realtime speed

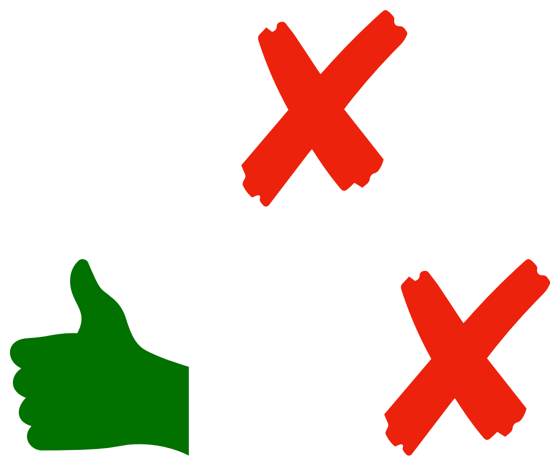
- Paper metrics:
  - mean average accuracy on camera pose

How the paper data & metric matches user?

Case study: image matching, metric

Note: not clear, why paper says that image matching is used in image retrieval
Add to the ToDo list to figure out (low priority)

- User metrics:
  - Image retrieval — Recall and precision
  - 3D reconstruction — many metrics
  - visual localization — camera pose accuracy
  - SLAM — camera pose accuracy, realtime speed



- Paper metrics:
  - mean average accuracy on camera pose

How the paper data & metric matches user?

Case study: image matching, metric

Note: not clear, why paper says that image matching is used in image retrieval
Add to the ToDo list to figure out (low priority)

- User metrics:
  - Image retrieval — Recall and precision
  - 3D reconstruction — many metrics
  - visual localization — camera pose accuracy
  - SLAM — camera pose accuracy, realtime speed

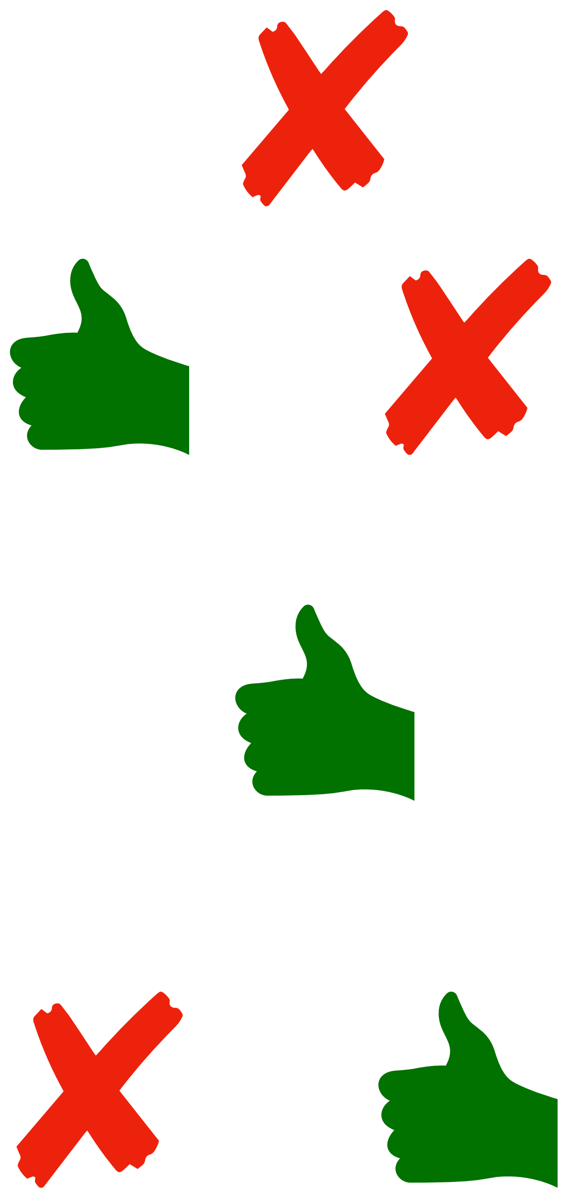
- Paper metrics:
  - mean average accuracy on camera pose

How the paper data & metric matches user?

Case study: image matching, metric

Note: not clear, why paper says that image matching is used in image retrieval
Add to the ToDo list to figure out (low priority)

- User metrics:
  - Image retrieval — Recall and precision
  - 3D reconstruction — many metrics
  - visual localization — camera pose accuracy
  - SLAM — camera pose accuracy, realtime speed



- Paper metrics:
  - mean average accuracy on camera pose

Case study: image matching

Case study: image matching, data

- Where the data comes from in production?
  - Image retrieval — user upload, can be anything
  - 3D reconstruction — professional cameras, drones, smartphone, historical photos, etc
  - visual localization — robot camera, car, smartphone, internet photo (OSI)
  - SLAM — robot, UAV, smart glasses, car

Case study: image matching

Case study: image matching, data

- Where the data comes from in production?
  - Image retrieval — user upload, can be anything
  - 3D reconstruction — professional cameras, drones, smartphone, historical photos, etc
  - visual localization — robot camera, car, smartphone, internet photo (OSI)
  - SLAM — robot, UAV, smart glasses, car

- Paper:
  - YCC100M dataset, from flickr (user photos)

Case study: image matching

Case study: image matching, data

- Where the data comes from in production?
  - Image retrieval — user upload, can be anything
  - 3D reconstruction — professional cameras, drones, smartphone, historical photos, etc
  - visual localization — robot camera, car, smartphone, internet photo (OSI)
  - SLAM — robot, UAV, smart glasses, car

- Paper:
  - YCC100M dataset, from flickr (user photos)

Case study: image matching

Case study: image matching, data

- Where the data comes from in production?
  - Image retrieval — user upload, can be anything
  - 3D reconstruction — professional cameras, drones, smartphone, historical photos, etc
  - visual localization — robot camera, car, smartphone, internet photo (OSI)
  - SLAM — robot, UAV, smart glasses, car

- Paper:
  - YCC100M dataset, from flickr (user photos)

Case study: image matching

Case study: image matching, data

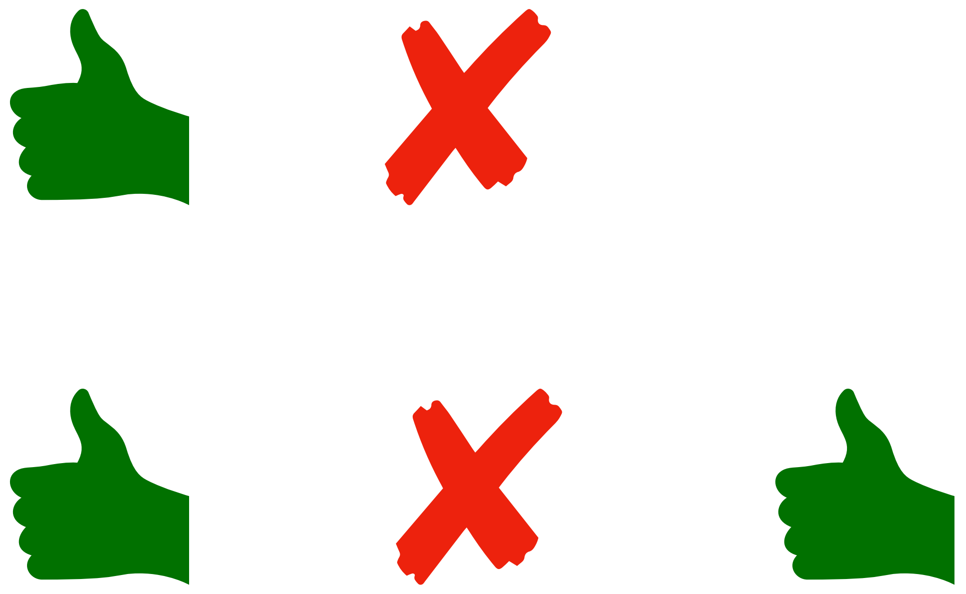
- Where the data comes from in production?
  - Image retrieval — user upload, can be anything
  - 3D reconstruction — professional cameras, drones, smartphone, historical photos, etc
  - visual localization — robot camera, car, smartphone, internet photo (OSI)
  - SLAM — robot, UAV, smart glasses, car

- Paper:
  - YCC100M dataset, from flickr (user photos)

Case study: image matching

Case study: image matching, data

- Where the data comes from in production?
  - Image retrieval — user upload, can be anything
  - 3D reconstruction — professional cameras, drones, smartphone, historical photos, etc
  - visual localization — robot camera, car, smartphone, internet photo (OSI)
  - SLAM — robot, UAV, smart glasses, car



- Paper:
  - YCC100M dataset, from flickr (user photos)

Case study: image matching

Case study: image matching, data

- Where the data comes from in production?
  - Image retrieval — user upload, can be anything
  - 3D reconstruction — professional cameras, drones, smartphone, historical photos, etc
  - visual localization — robot camera, car, smartphone, internet photo (OSI)
  - SLAM — robot, UAV, smart glasses, car

- Paper:
  - YCC100M dataset, from flickr (user photos)

Case study: image matching

Case study: image matching, data

- Where the data comes from in production?
  - Image retrieval — user upload, can be anything
  - 3D reconstruction — professional cameras, drones, smartphone, historical photos, etc
  - visual localization — robot camera, car, smartphone, internet photo (OSI)
  - SLAM — robot, UAV, smart glasses, car

- Paper:
  - YCC100M dataset, from flickr (user photos)

Case study: image matching

Case study: image matching, data

- Where the data comes from in production?
  - Image retrieval — user upload, can be anything 👍 ✗
  - 3D reconstruction — professional cameras, drones, smartphone, historical photos, etc 👍 ✗ 👍 ✗
  - visual localization — robot camera, car, smartphone, internet photo (OSI) ✗ ✗
  - SLAM — robot, UAV, smart glasses, car

- Paper:
  - YCC100M dataset, from flickr (user photos)

Case study: image matching

Case study: image matching, data

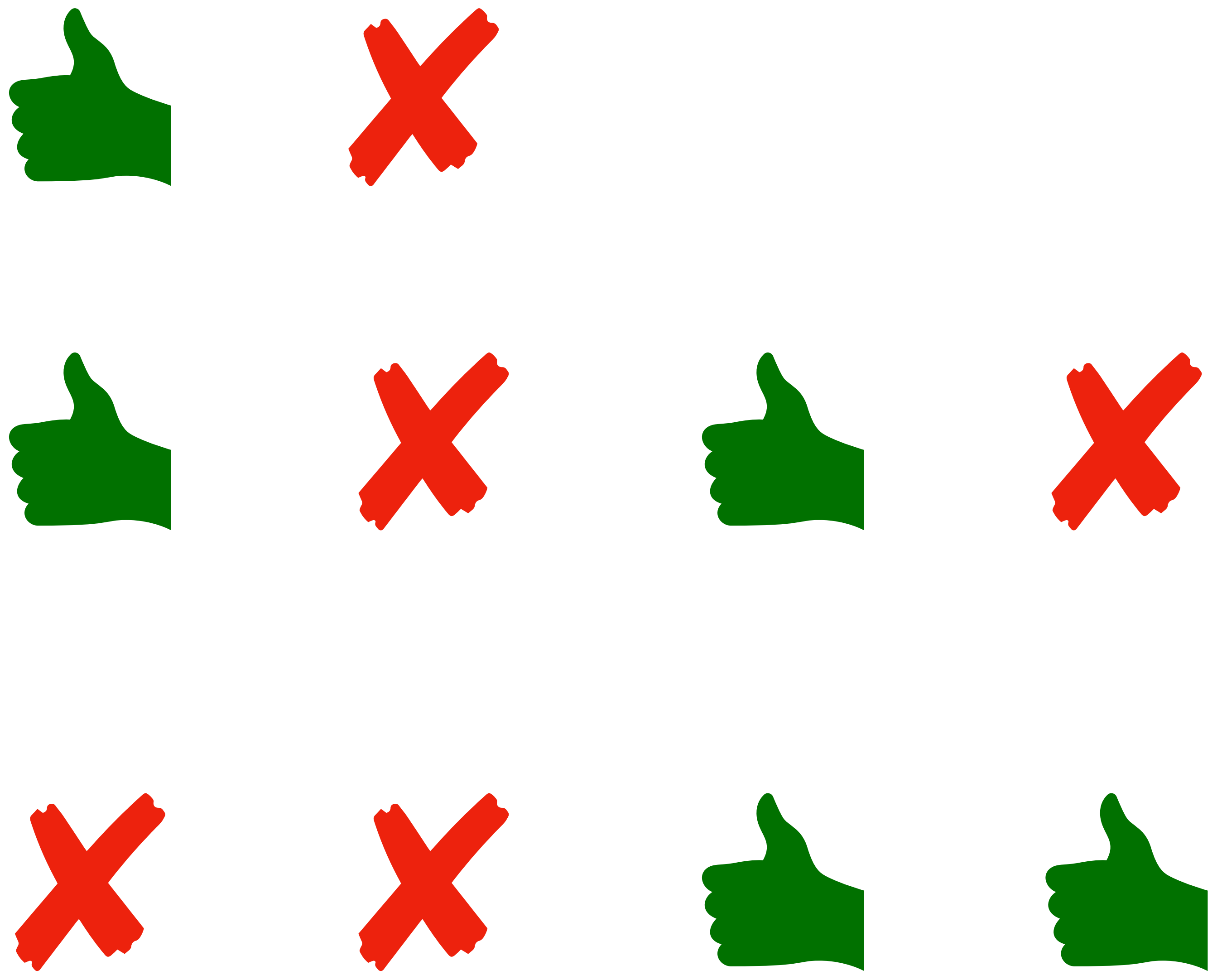
- Where the data comes from in production?
  - Image retrieval — user upload, can be anything
  - 3D reconstruction — professional cameras, drones, smartphone, historical photos, etc
  - visual localization — robot camera, car, smartphone, internet photo (OSI)
  - SLAM — robot, UAV, smart glasses, car

- Paper:
  - YCC100M dataset, from flickr (user photos)

Case study: image matching

Case study: image matching, data

- Where the data comes from in production?
  - Image retrieval — user upload, can be anything
  - 3D reconstruction — professional cameras, drones, smartphone, historical photos, etc
  - visual localization — robot camera, car, smartphone, internet photo (OSI)
  - SLAM — robot, UAV, smart glasses, car

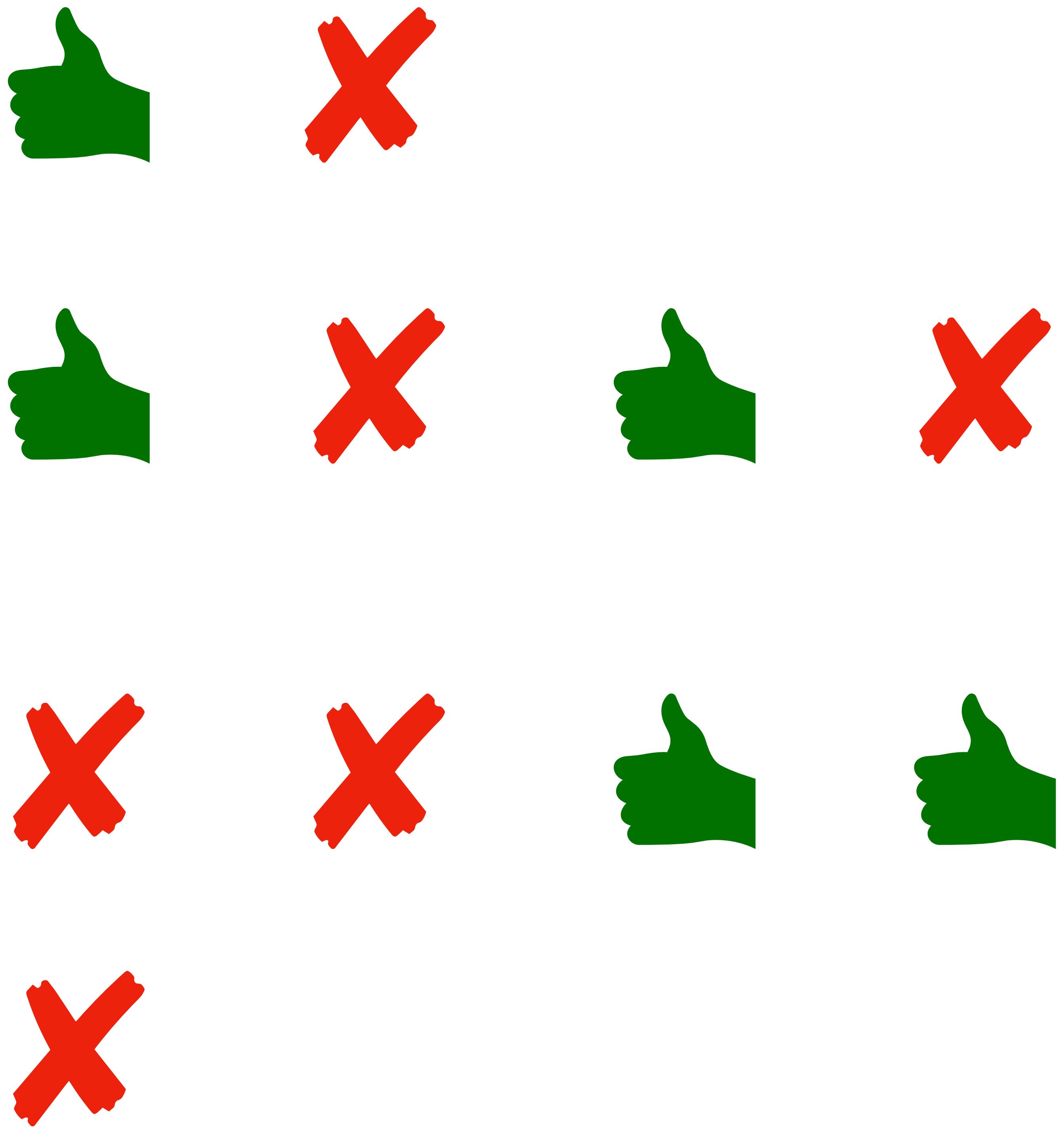


- Paper:
  - YCC100M dataset, from flickr (user photos)

Case study: image matching

Case study: image matching, data

- Where the data comes from in production?
  - Image retrieval — user upload, can be anything
  - 3D reconstruction — professional cameras, drones, smartphone, historical photos, etc
  - visual localization — robot camera, car, smartphone, internet photo (OSI)
  - SLAM — robot, UAV, smart glasses, car



- Paper:
  - YCC100M dataset, from flickr (user photos)

Case study: image matching

Case study: image matching, data

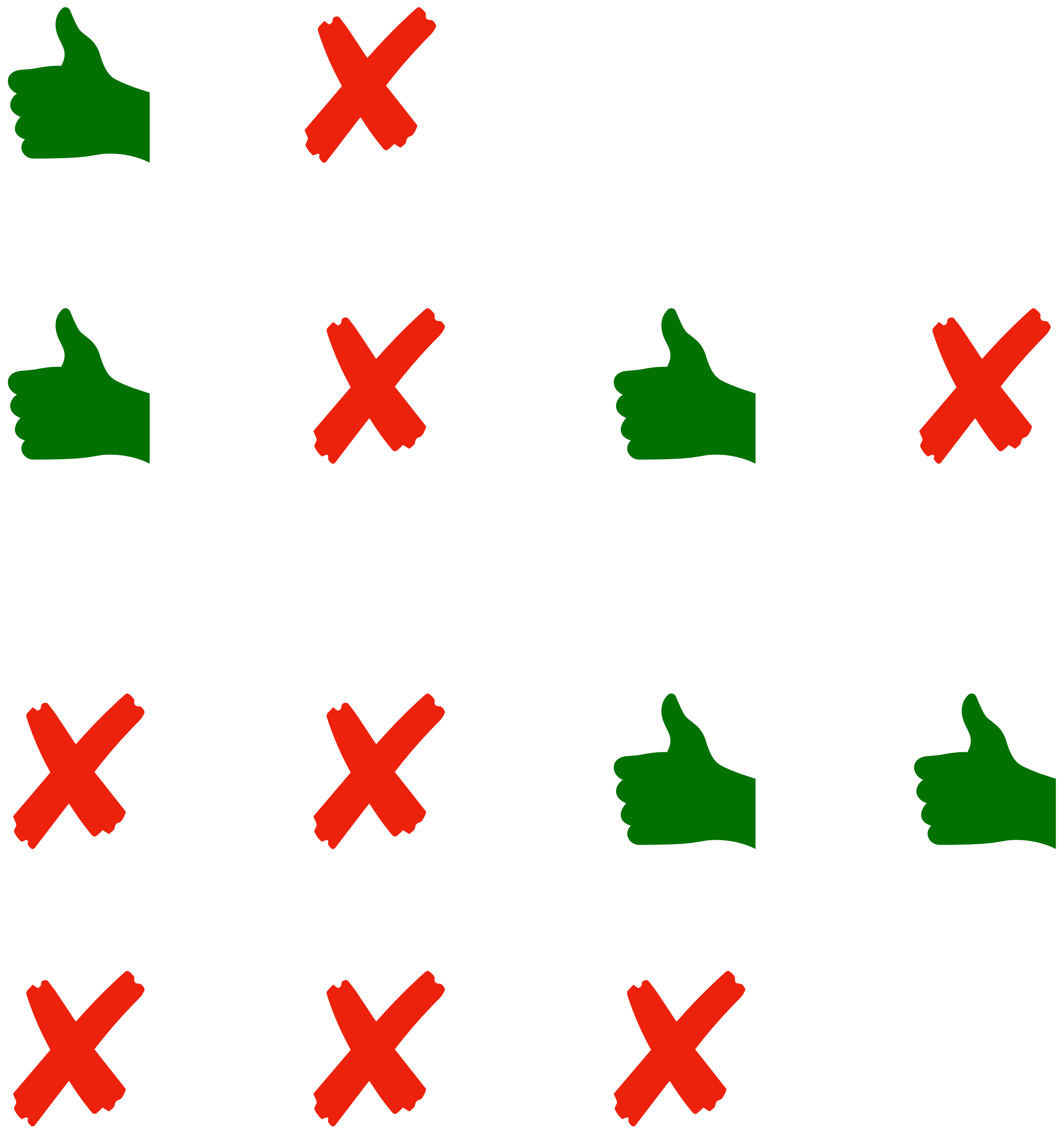
- Where the data comes from in production?
  - Image retrieval — user upload, can be anything
  - 3D reconstruction — professional cameras, drones, smartphone, historical photos, etc
  - visual localization — robot camera, car, smartphone, internet photo (OSI)
  - SLAM — robot, UAV, smart glasses, car

- Paper:
  - YCC100M dataset, from flickr (user photos)

Case study: image matching

Case study: image matching, data

- Where the data comes from in production?
  - Image retrieval — user upload, can be anything
  - 3D reconstruction — professional cameras, drones, smartphone, historical photos, etc
  - visual localization — robot camera, car, smartphone, internet photo (OSI)
  - SLAM — robot, UAV, smart glasses, car



- Paper:
  - YCC100M dataset, from flickr (user photos)

Case study: image matching

Case study: image matching, data

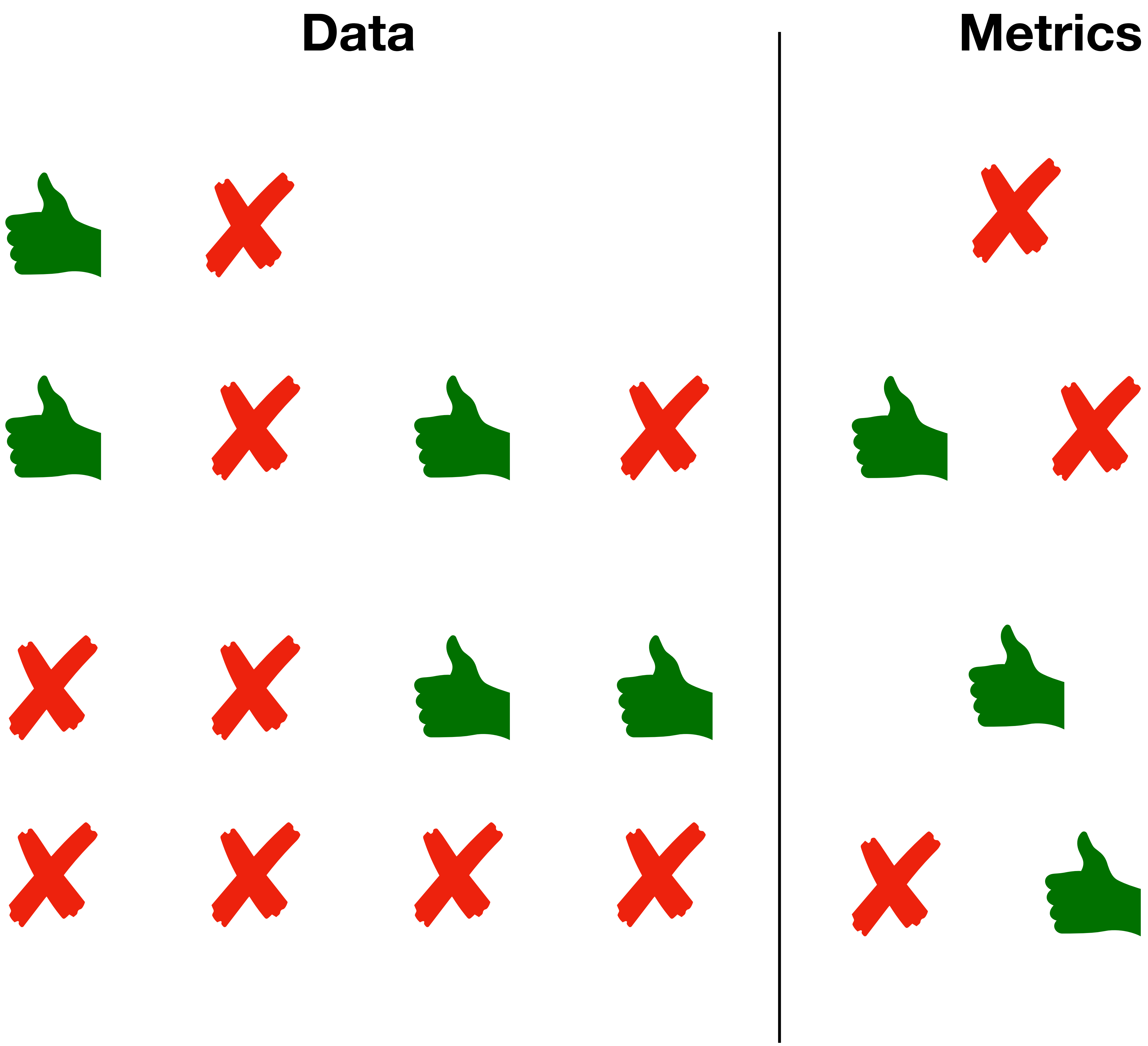
- Where the data comes from in production?
  - Image retrieval — user upload, can be anything 👍 ✗
  - 3D reconstruction — professional cameras, drones, smartphone, historical photos, etc 👍 ✗ 👍 ✗
  - visual localization — robot camera, car, smartphone, internet photo (OSI) ✗ ✗ 👍 👍
  - SLAM — robot, UAV, smart glasses, car ✗ ✗ ✗ ✗

- Paper:
  - YCC100M dataset, from flickr (user photos)

Case study: image matching

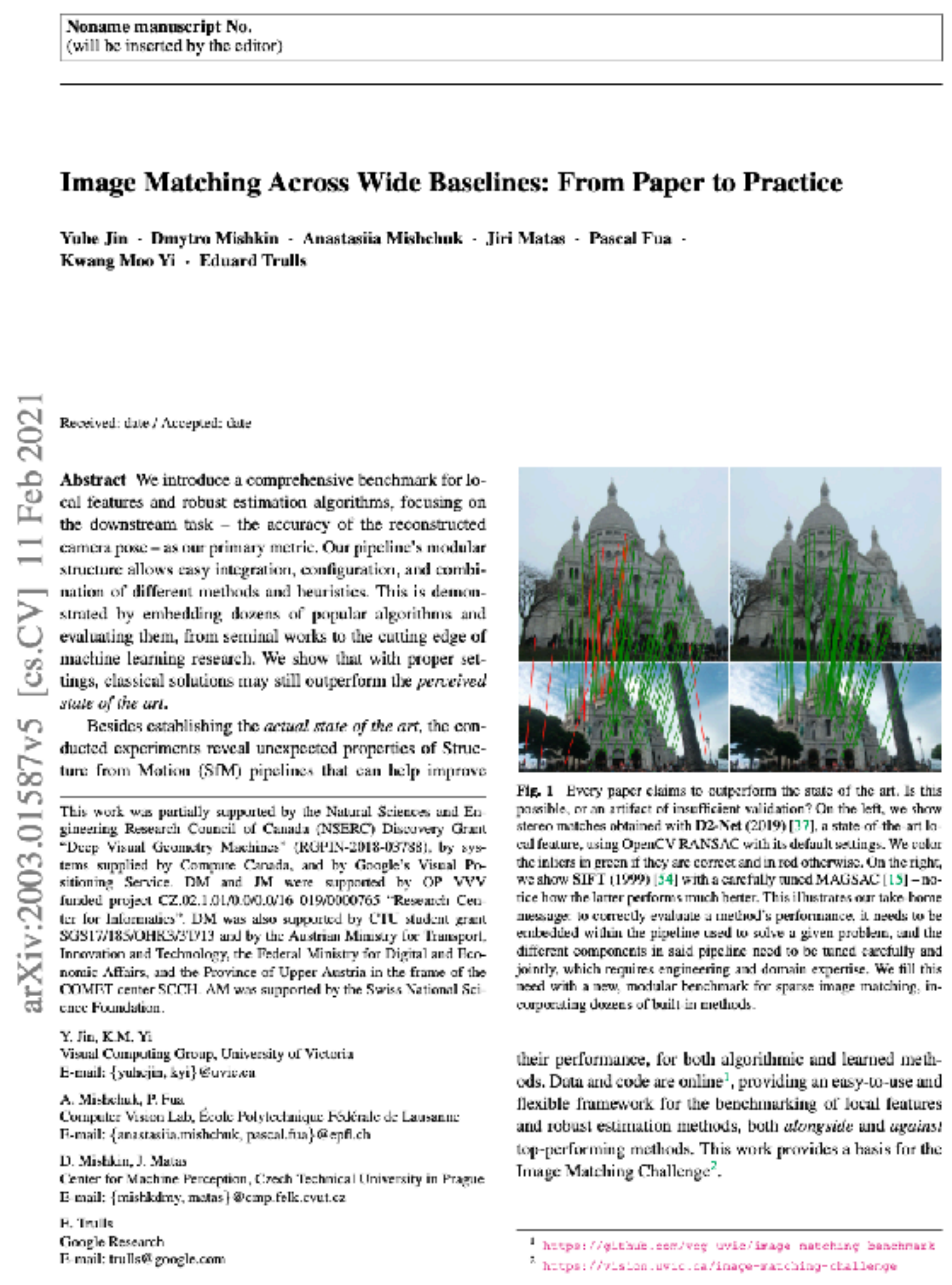
Case study: image matching, data

- How the paper results are relevant for
  - Image retrieval
  - 3D reconstruction
  - visual localization
  - SLAM



Intermediate conclusion

- The paper findings are partially relevant for some use-cases in
  - visual localization
  - 3D reconstruction
- Loosely relevant for the SLAM
- Not relevant at all for image retrieval



Noname manuscript No.
(will be inserted by the editor)

Image Matching Across Wide Baselines: From Paper to Practice

Yuhe Jin · Dmytro Mishkin · Anastasiia Mishchuk · Jiri Matas · Pascal Fua · Kwang Moo Yi · Eduard Trulls

arXiv:2003.01587v5 [cs.CV] 11 Feb 2021

Received: date / Accepted: date

Abstract We introduce a comprehensive benchmark for local features and robust estimation algorithms, focusing on the downstream task – the accuracy of the reconstructed camera pose – as our primary metric. Our pipeline's modular structure allows easy integration, configuration, and combination of different methods and heuristics. This is demonstrated by embedding dozens of popular algorithms and evaluating them, from seminal works to the cutting edge of machine learning research. We show that with proper settings, classical solutions may still outperform the *perceived state of the art*.

Besides establishing the *actual state of the art*, the conducted experiments reveal unexpected properties of Structure from Motion (SfM) pipelines that can help improve their performance, for both algorithmic and learned methods. Data and code are online[1], providing an easy-to-use and flexible framework for the benchmarking of local features and robust estimation methods, both *alongside* and *against* top-performing methods. This work provides a basis for the Image Matching Challenge[2].

Fig. 1 Every paper claims to outperform the state of the art. Is this possible, or an artifact of insufficient validation? On the left, we show stereo matches obtained with D2-Net (2019) [17], a state-of-the-art local feature, using OpenCV RANSAC with its default settings. We color the inliers in green if they are correct and in red otherwise. On the right, we show SIFT (1999) [54] with a carefully tuned MAGSAC [15] – notice how the latter performs much better. This illustrates our take-home message: to correctly evaluate a method's performance, it needs to be embedded within the pipeline used to solve a given problem, and the different components in said pipeline need to be tuned carefully and jointly, which requires engineering and domain expertise. We fill this need with a new, modular benchmark for sparse image matching, incorporating dozens of built-in methods.

This work was partially supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant "Deep Visual Geometry Machines" (RGPIN-2018-03788), by systems supplied by Compute Canada, and by Google's Visual Positioning Service. DM and JM were supported by OP VVV funded project CZ.02.1.01/0.0/0.0/16 019/0000765 "Research Center for Informatics". DM was also supported by CTU student grant SGS17/185/OHK3/3T/13 and by the Austrian Ministry for Transport, Innovation and Technology, the Federal Ministry for Digital and Economic Affairs, and the Province of Upper Austria in the frame of the COMET center SCCH. AM was supported by the Swiss National Science Foundation.

Y. Jin, K.M. Yi
Visual Computing Group, University of Victoria
E-mail: {yuhejin, kyi}@uvic.ca

A. Mishchuk, P. Fua
Computer Vision Lab, École Polytechnique Fédérale de Lausanne
E-mail: {anastasiia.mishchuk, pascal.fua}@epfl.ch

D. Mishkin, J. Matas
Center for Machine Perception, Czech Technical University in Prague
E-mail: {mishkdmy, matas}@cmp.felk.cvut.cz

E. Trulls
Google Research
E-mail: trulls@google.com

[1] https://github.com/vcg-uvic/image-matching-benchmark
[2] https://vision.uvic.ca/image-matching-challenge

Why should I do all this crap?

Why should I do all this crap?

Why should I do all this crap?

If you need a truck,
do not watch racing car reviews

Research sometimes deviates from practice

or often

https://twitter.com/ChristophMolnar/status/1485549716268109824

Research sometimes deviates from practice

or often

Christoph Molnar
@ChristophMolnar

A lot of machine learning research has detached itself from solving real problems, and created their own "benchmark-islands".

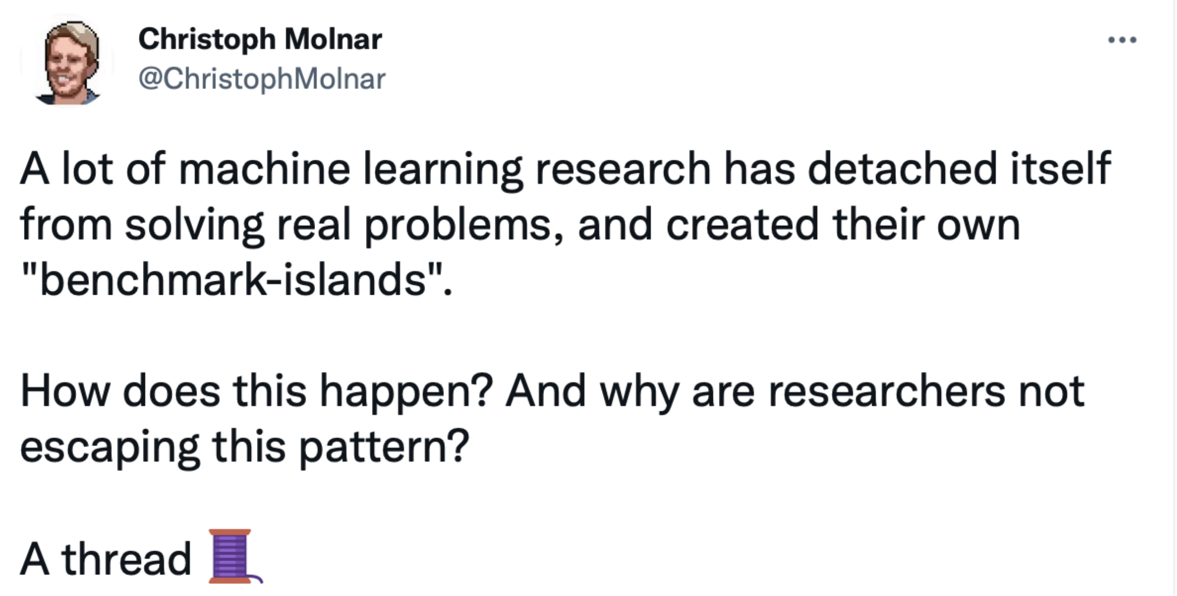
How does this happen? And why are researchers not escaping this pattern?

A thread 🧵

https://twitter.com/ChristophMolnar/status/1485549716268109824

Research sometimes deviates from practice

or often


Christoph Molnar
@ChristophMolnar

A lot of machine learning research has detached itself from solving real problems, and created their own "benchmark-islands".

How does this happen? And why are researchers not escaping this pattern?

A thread 🧵

https://twitter.com/ChristophMolnar/status/1485549716268109824

Research sometimes deviates from practice

or often

1. Establish a research topic. For example: Detect COVID-19 from x-ray images with machine learning.
2. First papers on this topic appear, and provide motivation for this research topic.
3. A precedent has been established: The initial papers pave the way for further research papers on the same topic.
4. The original motivation and assumptions for developing these predictive models are no longer questioned.
5. More researchers publish data and predictive models.

Christoph Molnar
@ChristophMolnar

A lot of machine learning research has detached itself from solving real problems, and created their own "benchmark-islands".

How does this happen? And why are researchers not escaping this pattern?

A thread 🧵

https://twitter.com/ChristophMolnar/status/1485549716268109824

Research sometimes deviates from practice

or often

Christoph Molnar
@ChristophMolnar

A lot of machine learning research has detached itself
from solving real problems, and created their own

s not

1. Establish a research topic. For example: Detect COVID-19 from x-ray images with machine learning.
2. First papers on this topic appear, and provide motivation for this research topic.
3. A precedent has been established: The initial papers pave the way for further research papers on the same topic.
4. The original motivation and assumptions for developing these predictive models are no longer questioned.
5. More researchers publish data and predictive models.
6. A community forms that cites and reviews each other.
7. Certain datasets become benchmarks and certain predictive models become state-of-the-art.
8. Predictive performance becomes the sole measure of progress, although improvements are becoming smaller.
9. Actual progress in solving the initial research questions has become irrelevant, even discouraged.
10. Decoupling from reality is complete.

https://twitter.com/ChristophMolnar/status/1485549716268109824

OK, paper is relevant

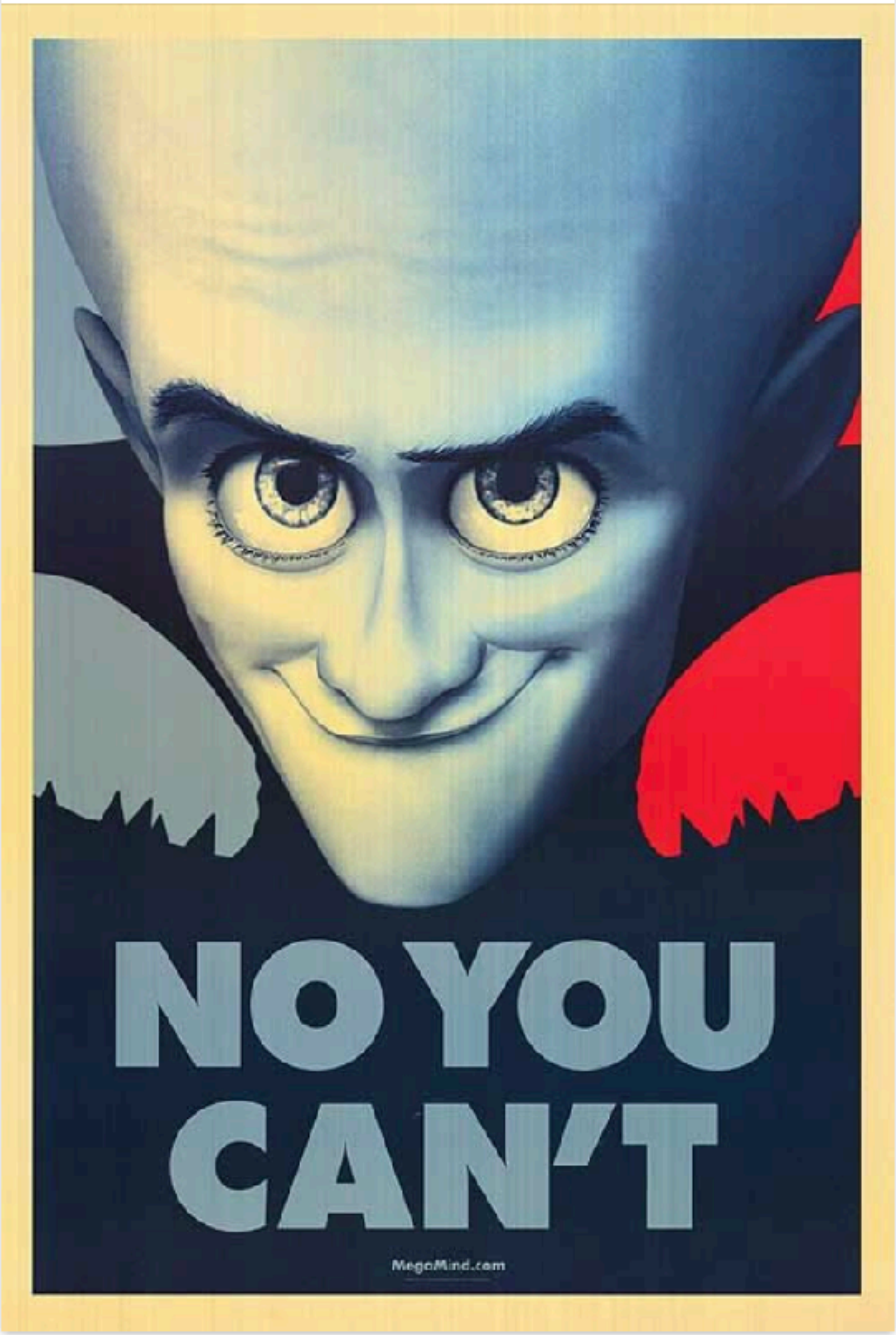
https://www.youtube.com/watch?v=_MeY1_ph2P4

OK, paper is relevant

https://www.youtube.com/watch?v=_MeY1_ph2P4

OK, paper is relevant

https://www.youtube.com/watch?v=_MeY1_ph2P4



Can we trust the paper?

"The only correct attitude, when reading the paper is:
Why do these lying bastard lie to me?"
David Forsyth

Do the metrics in the paper make sense?

If you don't have experience, you can skip this

Do the metrics in the paper make sense?

If you don't have experience, you can skip this

- Even if the metric is measuring a proper thing, it can still be misleading

Do the metrics in the paper make sense?

If you don't have experience, you can skip this

- Even if the metric is measuring a proper thing, it can still be misleading
  - Is it robust to outliers? E.g. 1 failure can kill average pose error over dataset

Do the metrics in the paper make sense?

If you don't have experience, you can skip this

- Even if the metric is measuring a proper thing, it can still be misleading
  - Is it robust to outliers? E.g. 1 failure can kill average pose error over dataset
  - Isn't it too robust to outliers? E.g. median error does not change if half of dataset is not solved at all

Do the metrics in the paper make sense?

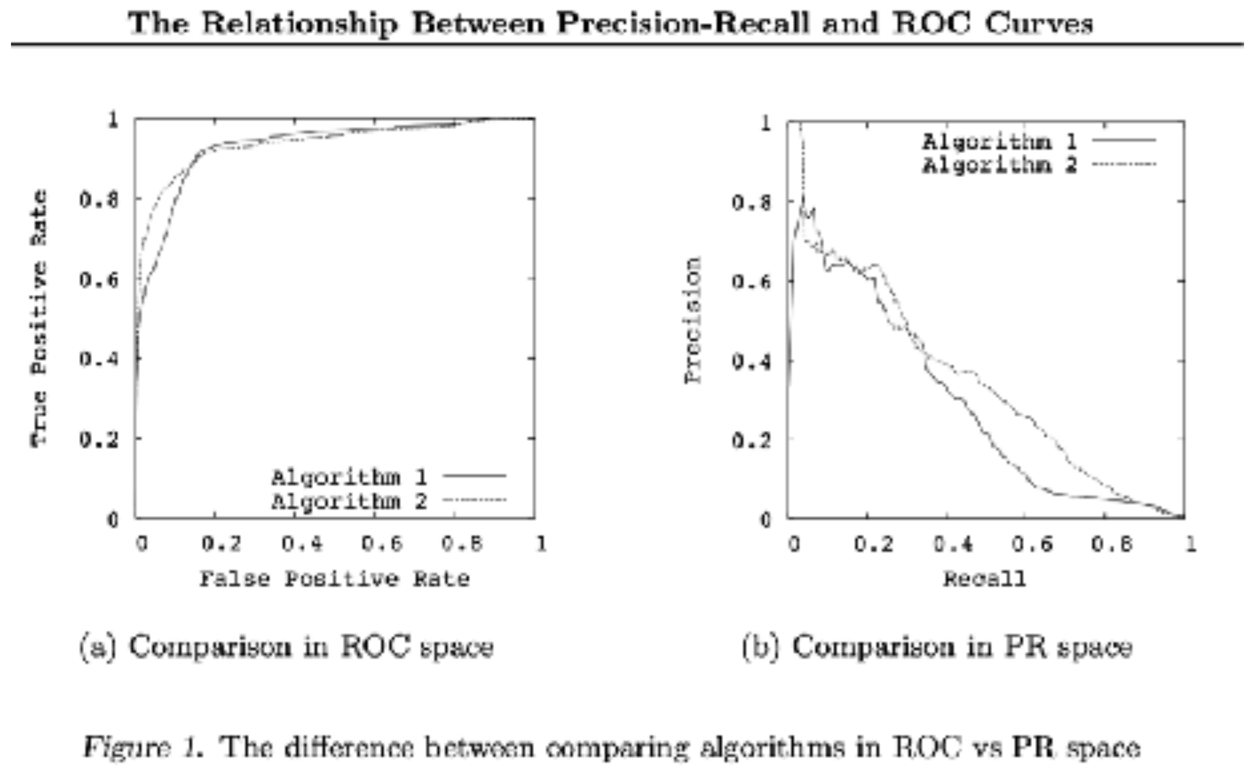
If you don't have experience, you can skip this

- Even if the metric is measuring a proper thing, it can still be misleading
  - Is it robust to outliers? E.g. 1 failure can kill average pose error over dataset
  - Isn't it too robust to outliers? E.g. median error does not change if half of dataset is not solved at all
  - What is the "null" performance according to the metric?

Do the metrics in the paper make sense?

If you don't have experience, you can skip this

- Even if the metric is measuring a proper thing, it can still be misleading
  - Is it robust to outliers? E.g. 1 failure can kill average pose error over dataset
  - Isn't it too robust to outliers? E.g. median error does not change if half of dataset is not solved at all
  - What is the "null" performance according to the metric?



https://www.biostat.wisc.edu/~page/rocpr.pdf

How the ground truth labels were obtained?

- Check if the annotation process makes sense
- Is it biased?

Noname manuscript No.
(will be inserted by the editor)

ImageNet Large Scale Visual Recognition Challenge

Olga Russakovsky* · Jia Deng* · Hao Su · Jonathan Krause · Sanjeev Satheesh · Sean Ma · Zhiheng Huang · Andrej Karpathy · Aditya Khosla · Michael Bernstein · Alexander C. Berg · Li Fei-Fei

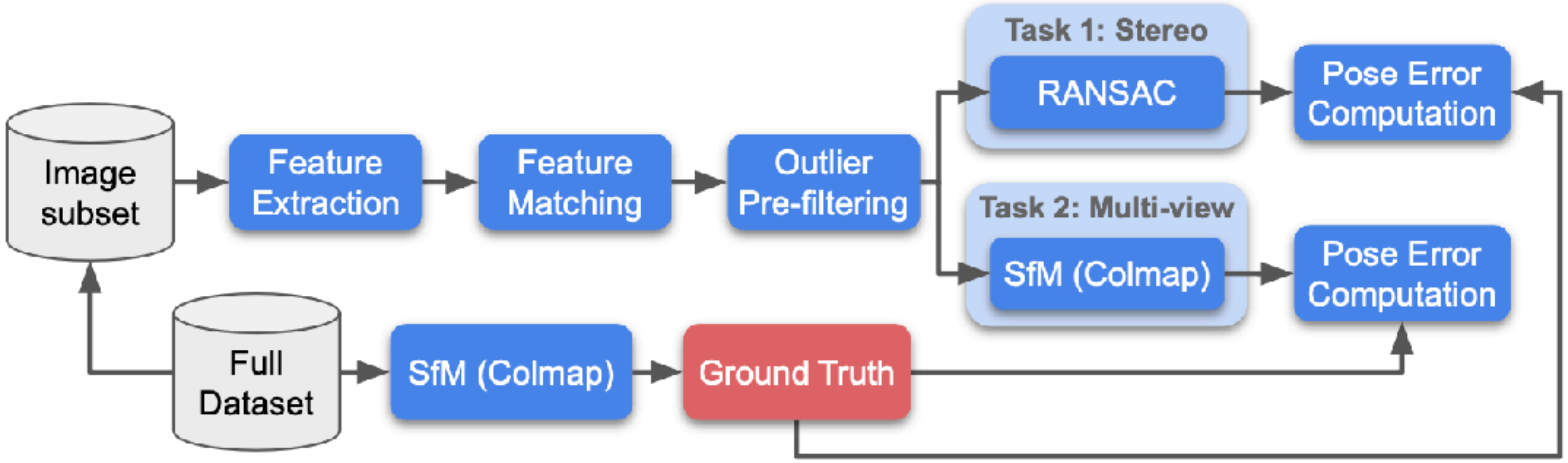
1. **Drawing** A worker draws one bounding box around one instance of an object on the given image.
2. **Quality verification** A second worker checks if the bounding box is correctly drawn.
3. **Coverage verification** A third worker checks if all object instances have bounding boxes.

https://arxiv.org/pdf/1409.0575.pdf

How the ground truth labels were obtained?

Case study: image matching

- Assumption: COLMAP reconstruction from the 1000 images outputs poses, which are good enough to evaluate stereo poses



How the ground truth labels were obtained?

Case study: image matching

- Assumption: COLMAP reconstruction from the 1000 images outputs poses, which are good enough to evaluate stereo poses
- Is it validated?

3.3 On the quality of the "ground-truth"

Our core assumption is that accurate poses can be obtained from large sets of images without human intervention. Such poses are used as the "ground truth" for evaluation of image matching performance on pairs or small subsets of images – a harder, proxy task. Should this assumption hold, the (relative) poses retrieved with a large enough number of images would not change as more images are added, and these poses would be the same regardless of which local feature is used. To validate this, we pick the scene "Sacre Coeur" and compute SfM reconstructions with a varying number of images: 100, 200, 400, 800, and 1179 images (the entire "Sacre Coeur" dataset), where each set contains the previous one; new images are being added and no images are removed. We run each reconstruction three times, and report the aver-

https://arxiv.org/pdf/1409.0575.pdf

How the ground truth labels were obtained?

Case study: image matching

- Is assumption validated?

https://arxiv.org/pdf/1409.0575.pdf

How the ground truth labels were obtained?

Case study: image matching

- Is assumption validated?

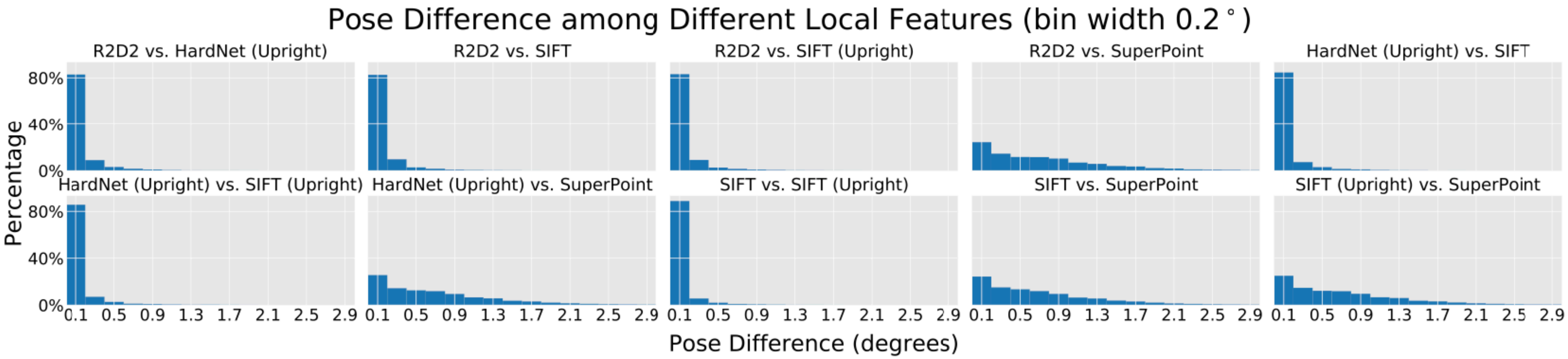


Fig. 7 Histograms of pose differences between reconstructions with different local feature methods. We consider five different local features – including rotation-sensitive and upright SIFT – resulting in 10 combinations. The plots show that about 80% percent of image pairs are within a 0.2^o pose difference, with the exception of those involving SuperPoint.

https://arxiv.org/pdf/1409.0575.pdf

How the ground truth labels were obtained?

Case study: image matching

- Is assumption validated?

https://arxiv.org/pdf/1409.0575.pdf

How the ground truth labels were obtained?

Case study: image matching

- Is assumption validated?

Local featured type	Number of Images				
	100	200	400	800	all
SIFT [54]	0.06°	0.09°	0.06°	0.07°	0.09°
SIFT (Upright) [54]	0.07°	0.07°	0.04°	0.06°	0.09°
HardNet (Upright) [62]	0.06°	0.06°	0.06°	0.04°	0.05°
SuperPoint [34]	0.31°	0.25°	0.33°	0.19°	0.32°
R2D2 [80]	0.12°	0.08°	0.07°	0.08°	0.05°

Table 2 Standard deviation of the pose difference of three COLMAP runs with different number of images. Most of them are below 0.1°, except for SuperPoint.

https://arxiv.org/pdf/1409.0575.pdf

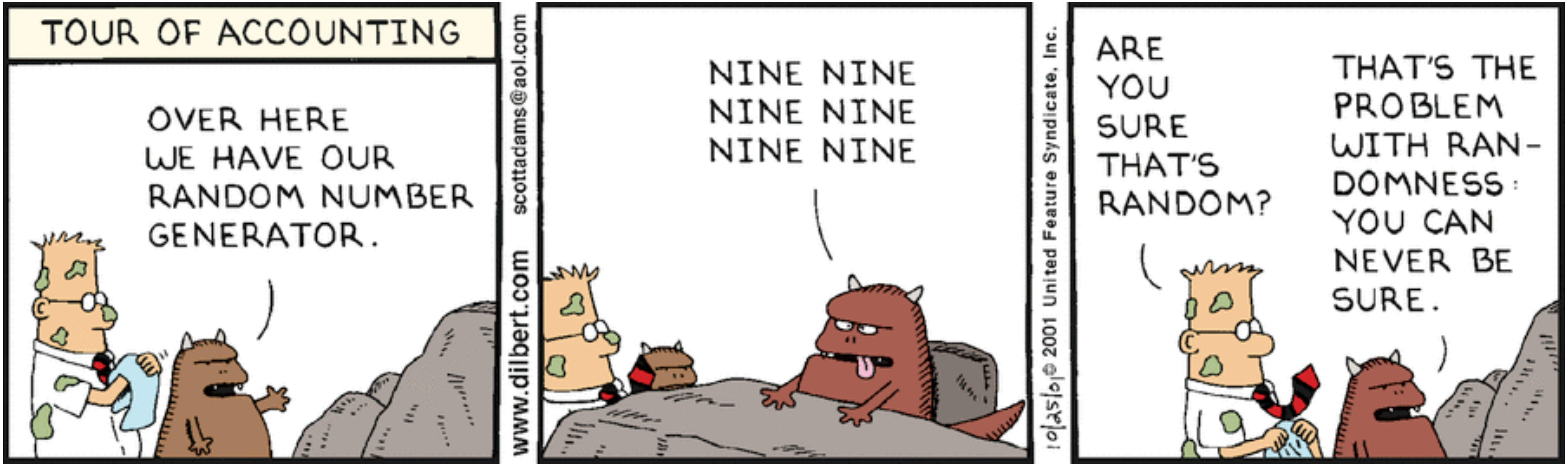
It seems that we can trust the paper…

It seems that we can trust the paper…

- Although we never can be sure

It seems that we can trust the paper…

- Although we never can be sure



You should never be 100% sure

I am not joking

You should never be 100% sure

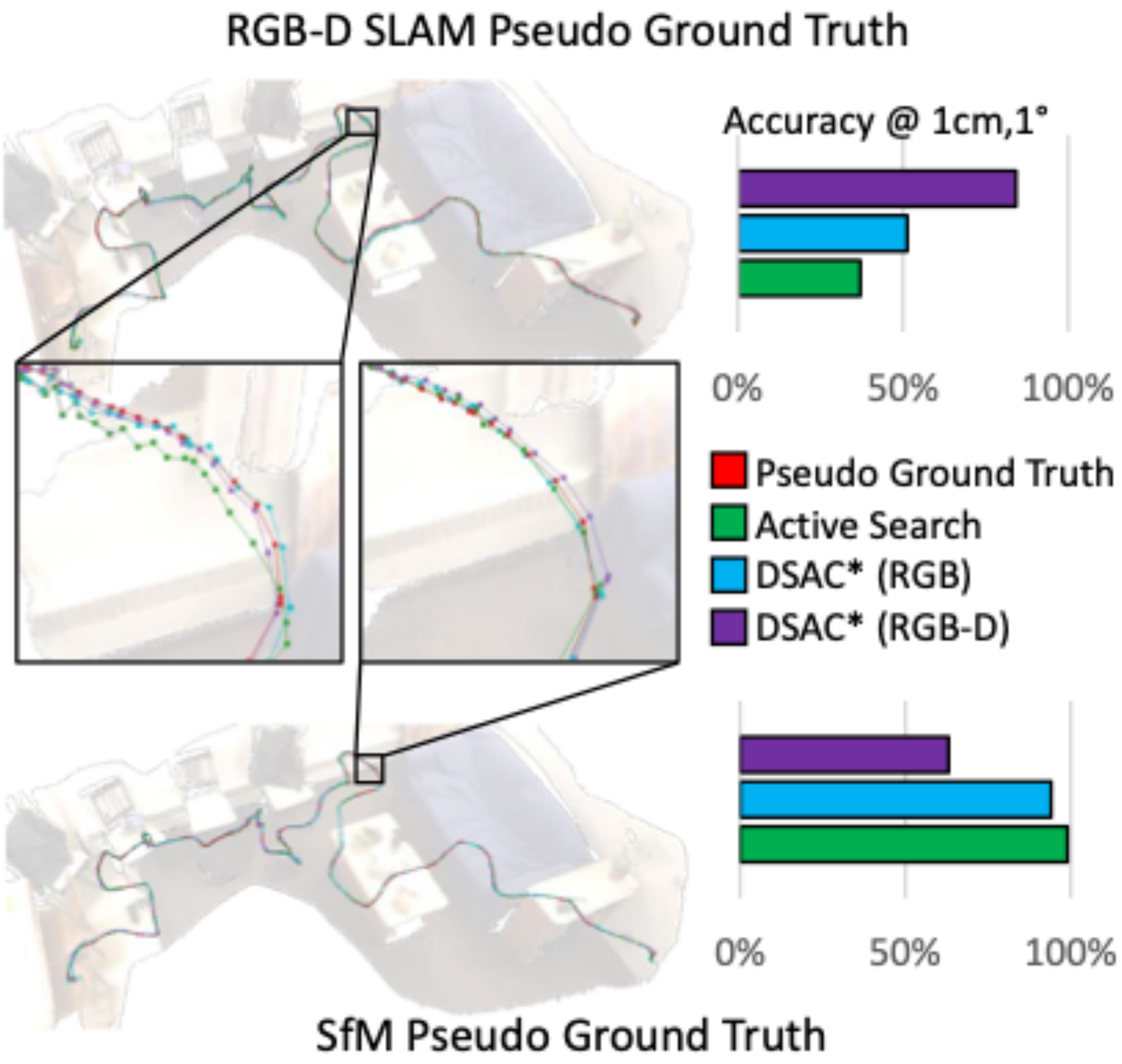
I am not joking

- On the Limits of Pseudo Ground Truth in Visual Camera Re-localisation. Brachmann et.al ICCV 2021

You should never be 100% sure

I am not joking

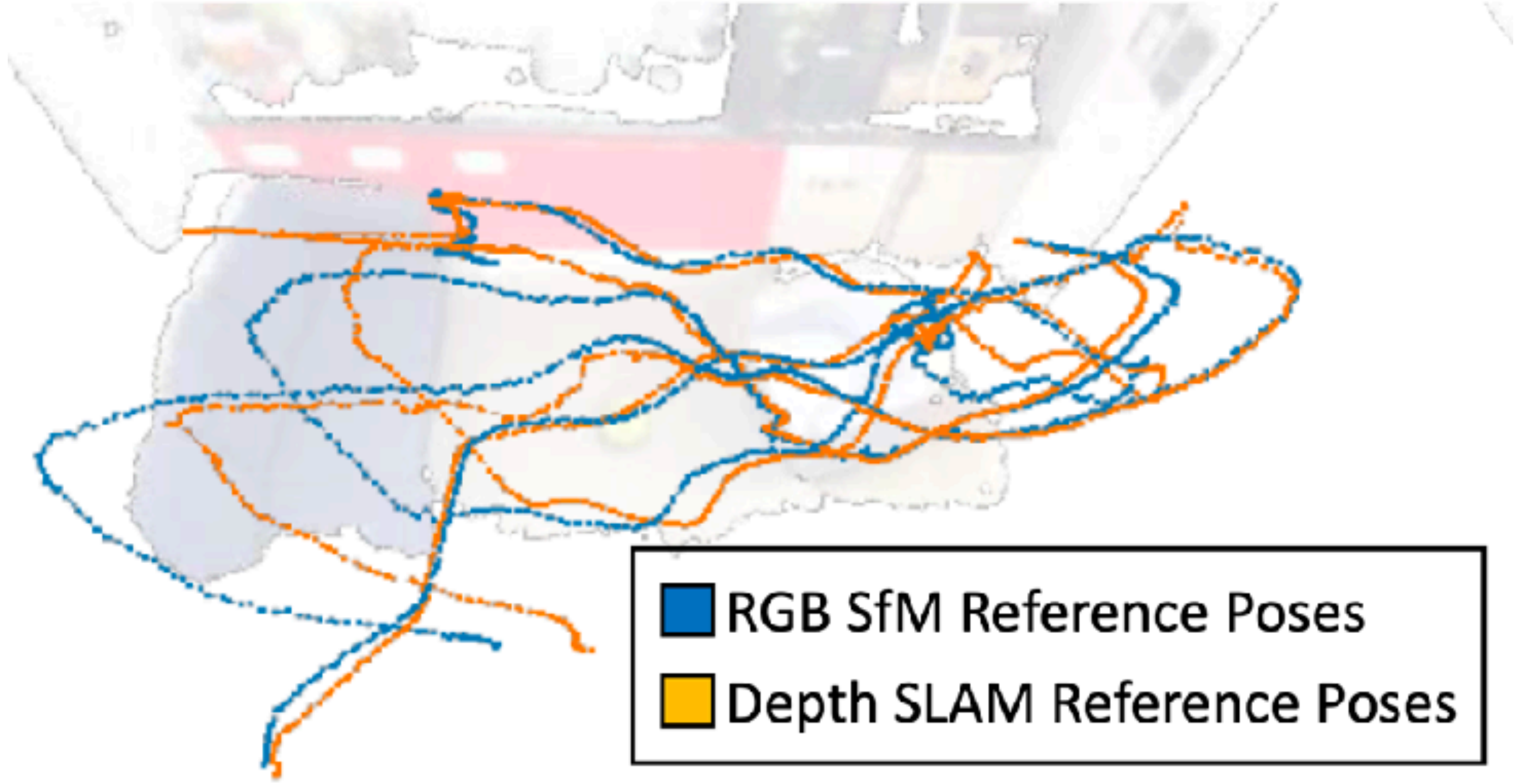
- On the Limits of Pseudo Ground Truth in Visual Camera Re-localisation. Brachmann et.al ICCV 2021

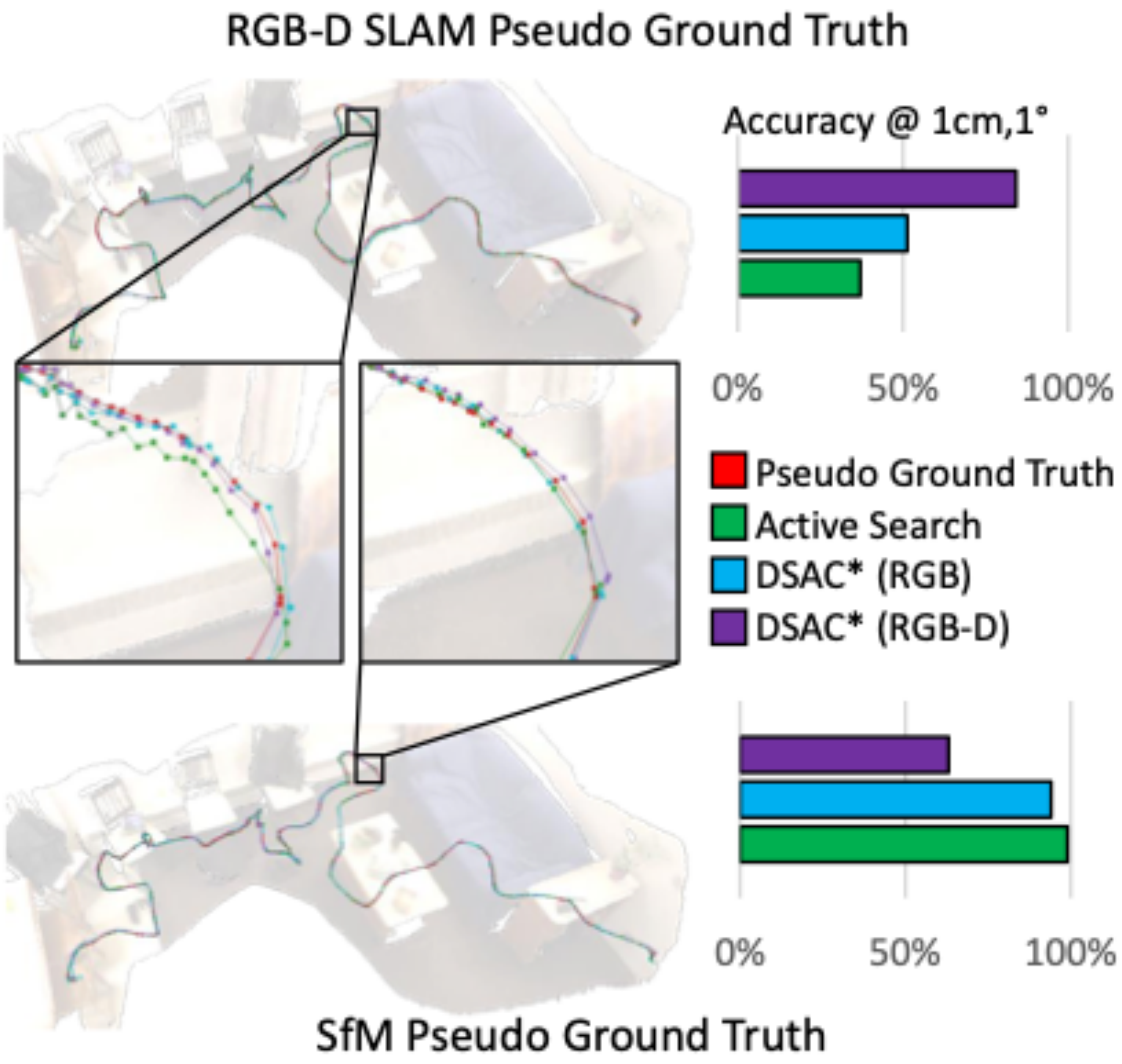


You should never be 100% sure

I am not joking

- On the Limits of Pseudo Ground Truth in Visual Camera Re-localisation. Brachmann et.al ICCV 2021

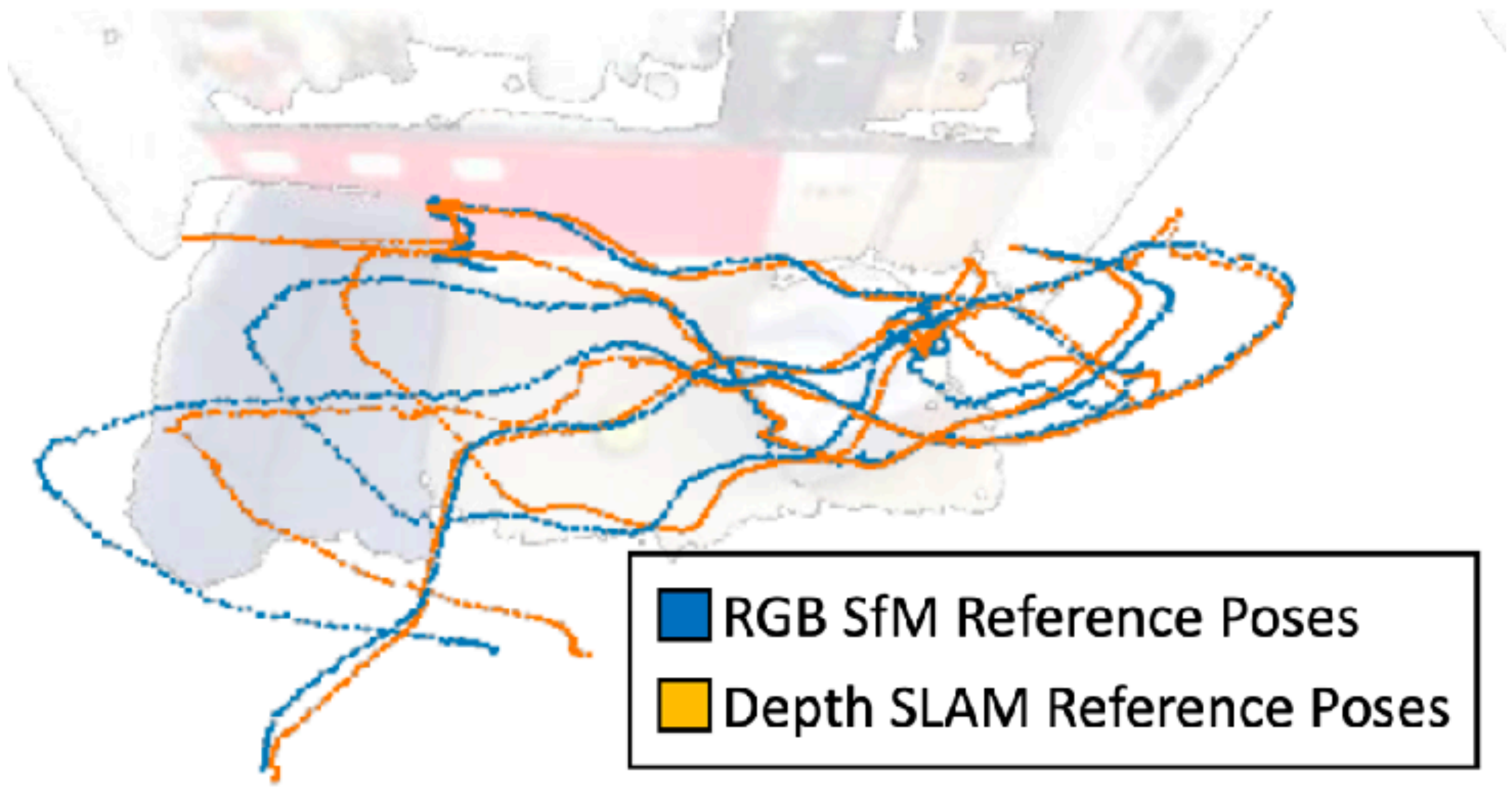


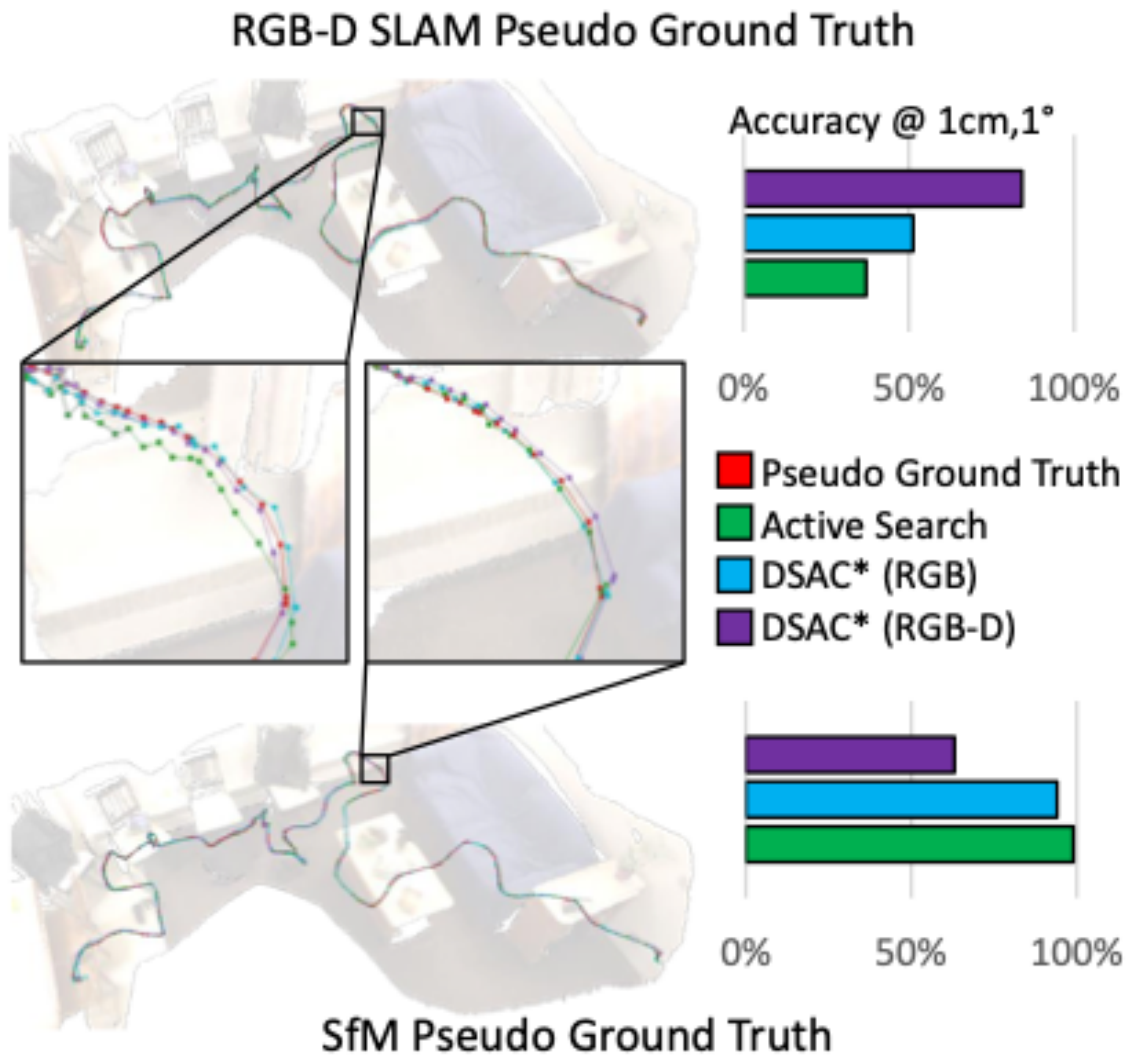


You should never be 100% sure

I am not joking

- On the Limits of Pseudo Ground Truth in Visual Camera Re-localisation. Brachmann et.al ICCV 2021





https://ducha-aiki.github.io/wide-baseline-stereo-blog/2021/09/04/review-of-pseudoGT-limitations.html

Paper is relevant and seems truthworthy

https://www.youtube.com/watch?v=_MeY1_ph2P4

Source

Paper is relevant and seems truthworthy

https://www.youtube.com/watch?v=_MeY1_ph2P4

Source

Paper is relevant and seems truthworthy

https://www.youtube.com/watch?v=_MeY1_ph2P4

Source

Checking results and methods

Start with baseline

Checking results and methods

Start with baseline

- Find the method, which is the established “gold standard” and/or the baseline

Checking results and methods

Start with baseline

- Find the method, which is the established "gold standard" and/or the baseline

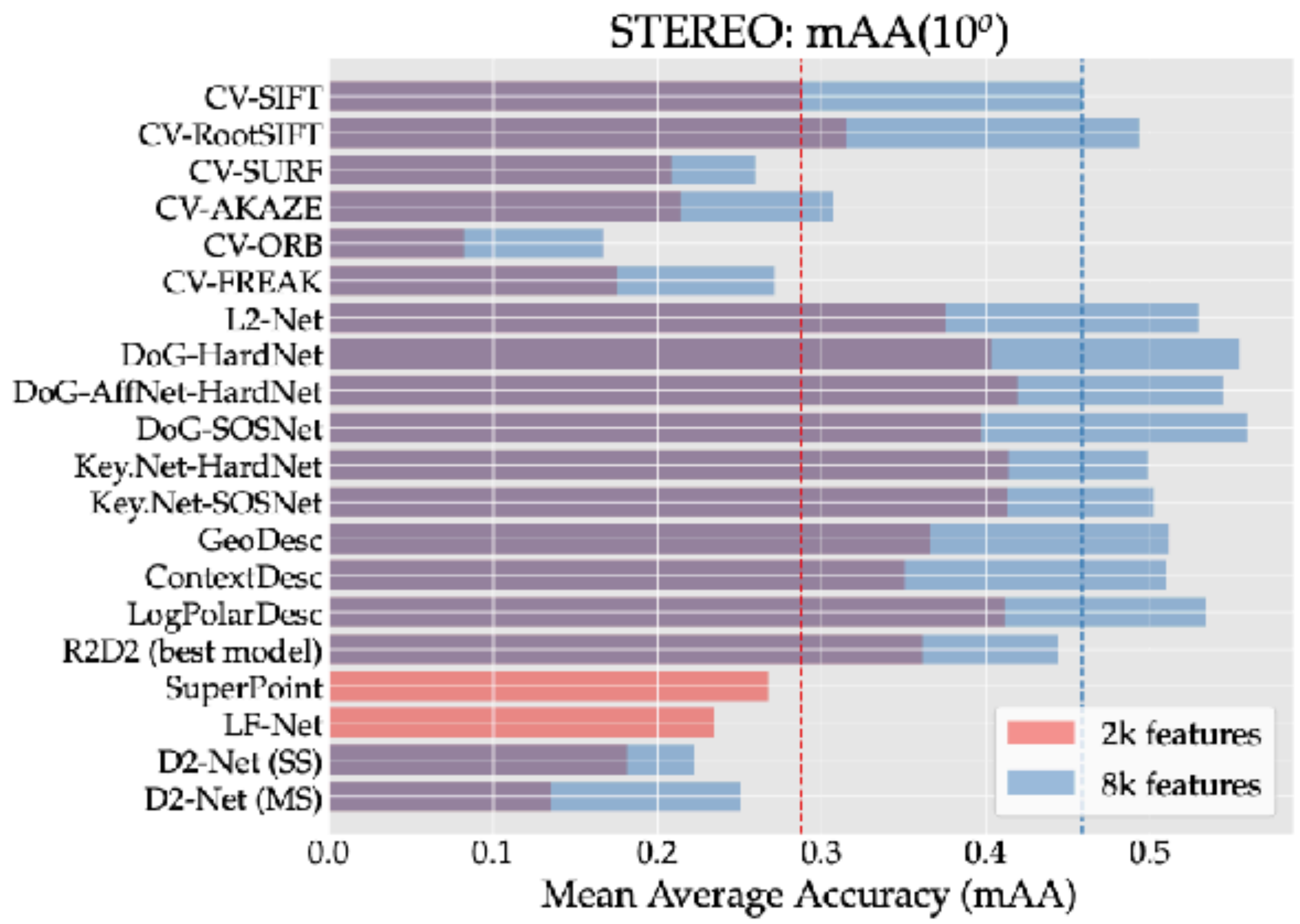


Fig. 17 Test – Stereo performance: 2k vs 8k features. We compare the results obtained with different methods using either 2k or 8k features – we use DEGENSAC, which performs better than other RANSAC variants under most circumstances. Dashed lines indicate SIFT's performance. For LF-Net and SuperPoint we do not include results with 8k features, as we failed to obtain meaningful results. For R2D2, we use the best model for each setting.

Checking results and methods

Start with baseline

- Find the method, which is the established “gold standard” and/or the baseline
- For image matching it is SIFT

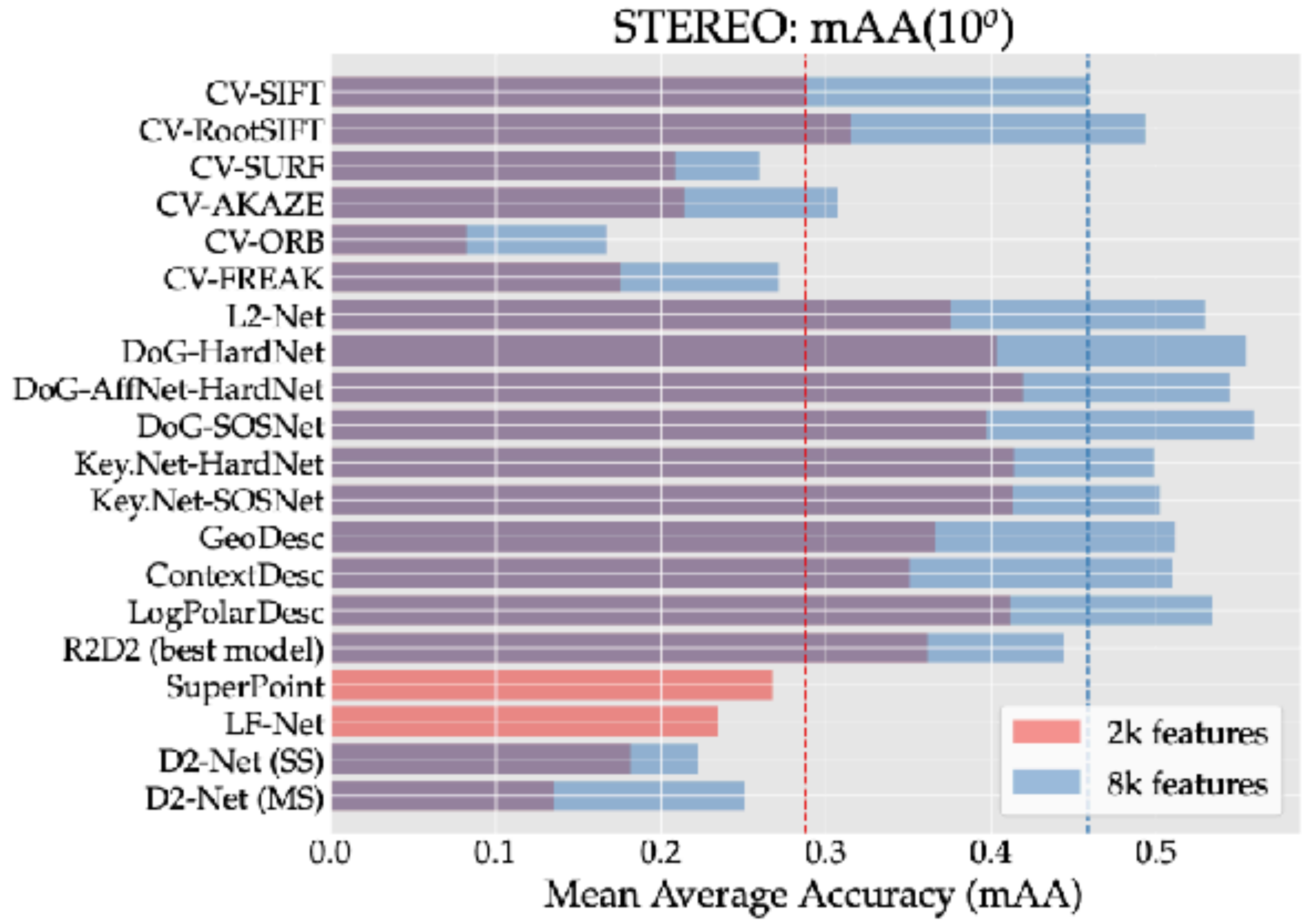


Fig. 17 Test – Stereo performance: 2k vs 8k features. We compare the results obtained with different methods using either 2k or 8k features – we use DEGENSAC, which performs better than other RANSAC variants under most circumstances. Dashed lines indicate SIFT’s performance. For LF-Net and SuperPoint we do not include results with 8k features, as we failed to obtain meaningful results. For R2D2, we use the best model for each setting.

Checking results and methods

Start with baseline

- Find the method, which is the established “gold standard” and/or the baseline
- For image matching it is SIFT
- Check how it works

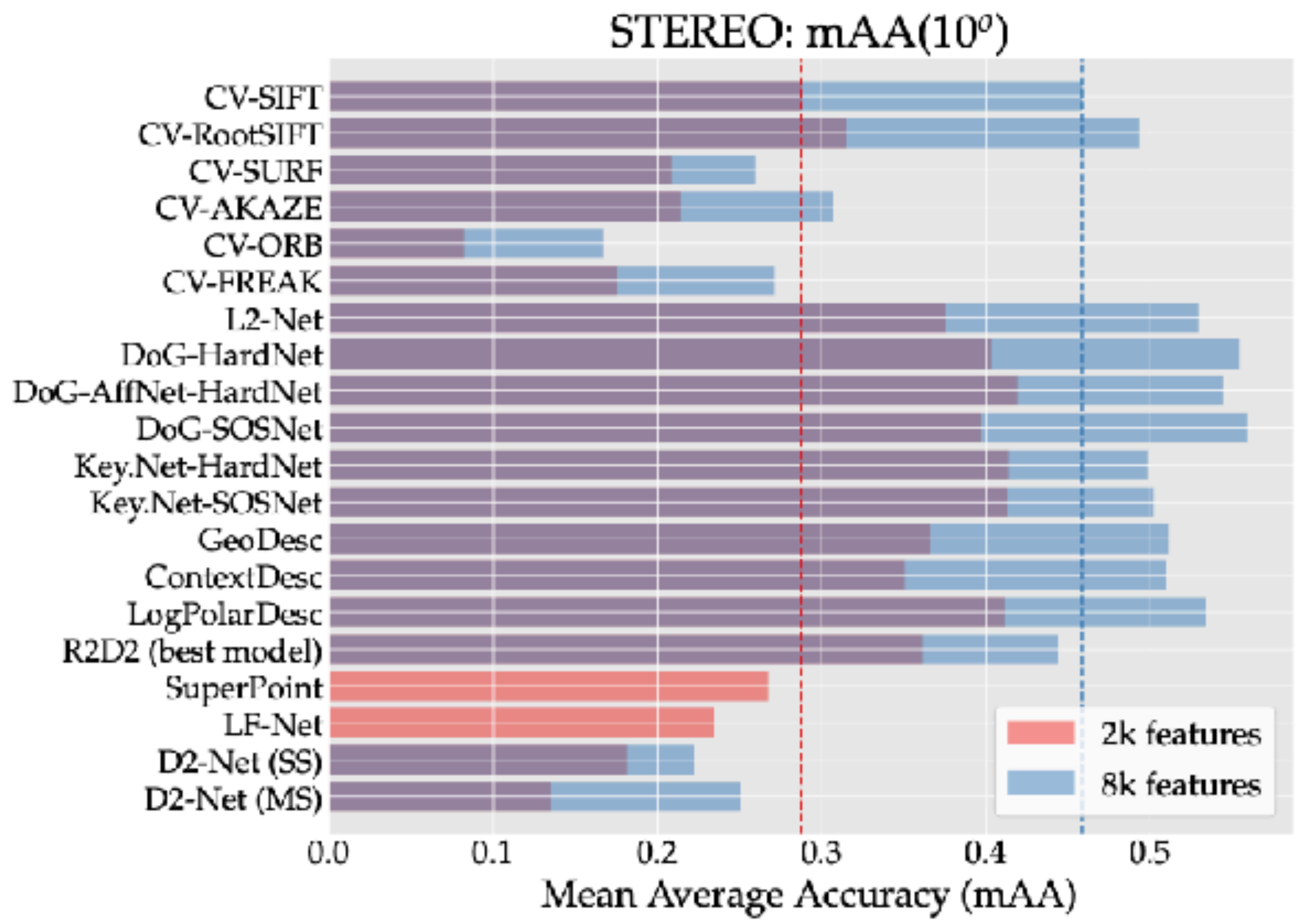


Fig. 17 Test – Stereo performance: 2k vs 8k features. We compare the results obtained with different methods using either 2k or 8k features – we use DEGENSAC, which performs better than other RANSAC variants under most circumstances. Dashed lines indicate SIFT’s performance. For LF-Net and SuperPoint we do not include results with 8k features, as we failed to obtain meaningful results. For R2D2, we use the best model for each setting.

Checking results and methods

Start with baseline

- Find the method, which is the established “gold standard” and/or the baseline
- For image matching it is SIFT
- Check how it works
  - Note, that the “gold standard” might be complex and highly engineered

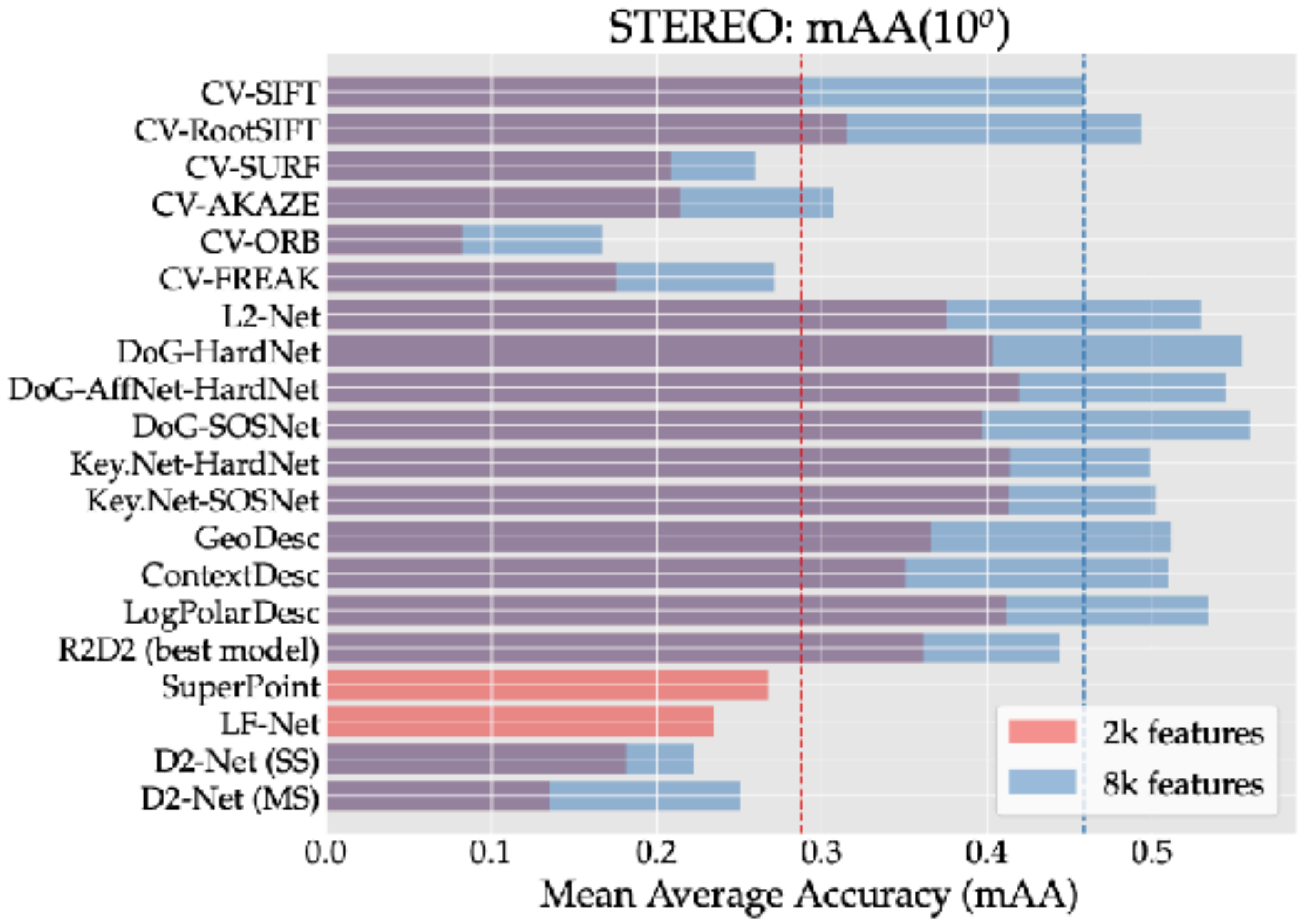


Fig. 17 Test – Stereo performance: 2k vs 8k features. We compare the results obtained with different methods using either 2k or 8k features – we use DEGENSAC, which performs better than other RANSAC variants under most circumstances. Dashed lines indicate SIFT’s performance. For LF-Net and SuperPoint we do not include results with 8k features, as we failed to obtain meaningful results. For R2D2, we use the best model for each setting.

Checking results and methods

Start with baseline

- Find the method, which is the established "gold standard" and/or the baseline
- For image matching it is SIFT
- Check how it works
  - Note, that the "gold standard" might be complex and highly engineered
  - New methods might be simpler to implement

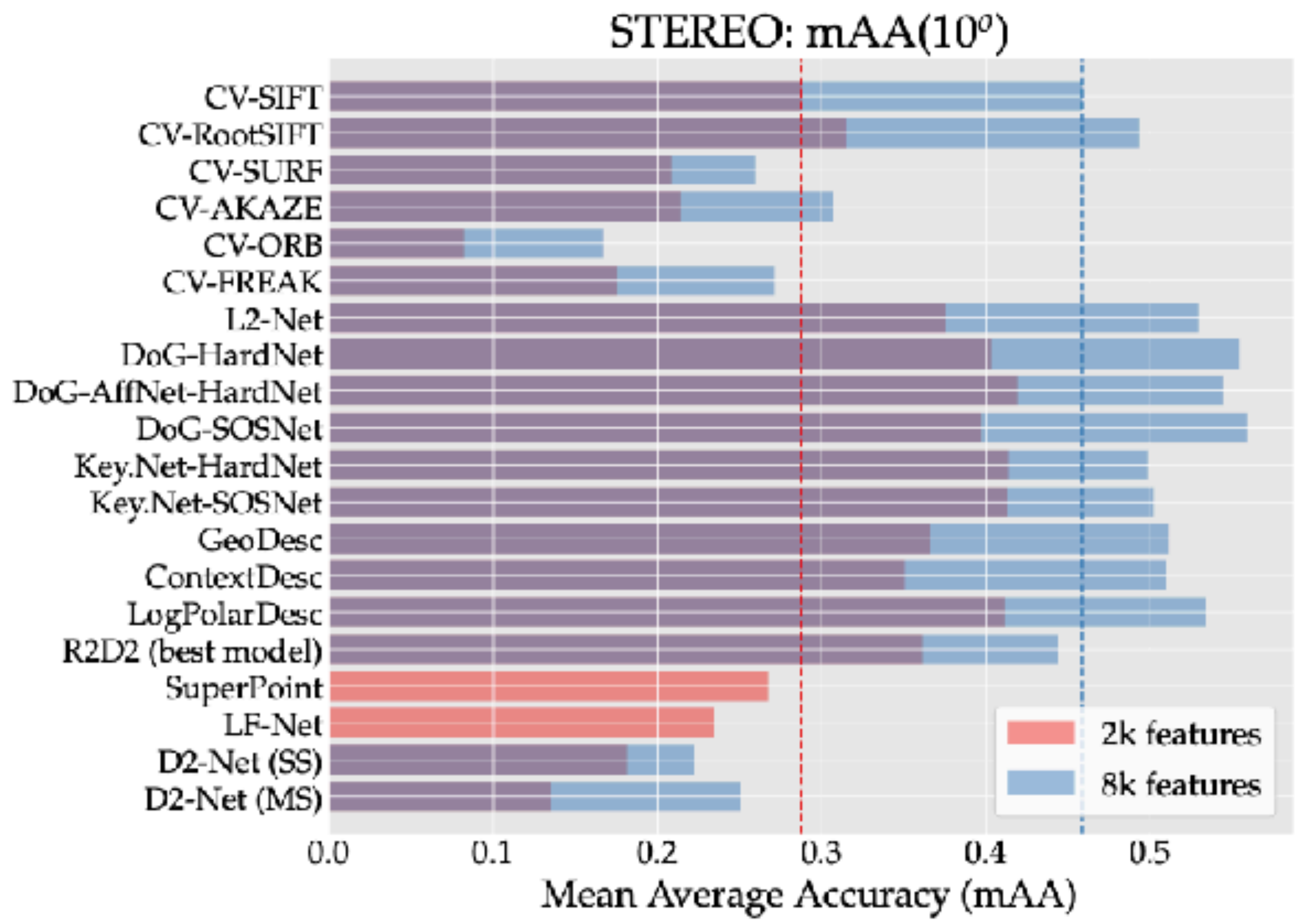


Fig. 17 Test – Stereo performance: 2k vs 8k features. We compare the results obtained with different methods using either 2k or 8k features – we use DEGENSAC, which performs better than other RANSAC variants under most circumstances. Dashed lines indicate SIFT's performance. For LF-Net and SuperPoint we do not include results with 8k features, as we failed to obtain meaningful results. For R2D2, we use the best model for each setting.

Checking results and methods

Start with baseline

- Find the method, which is the established “gold standard” and/or the baseline
- For image matching it is SIFT
- Check how it works
    - Note, that the “gold standard” might be complex and highly engineered
    - New methods might be simpler to implement
- Play selected (good and simple) methods

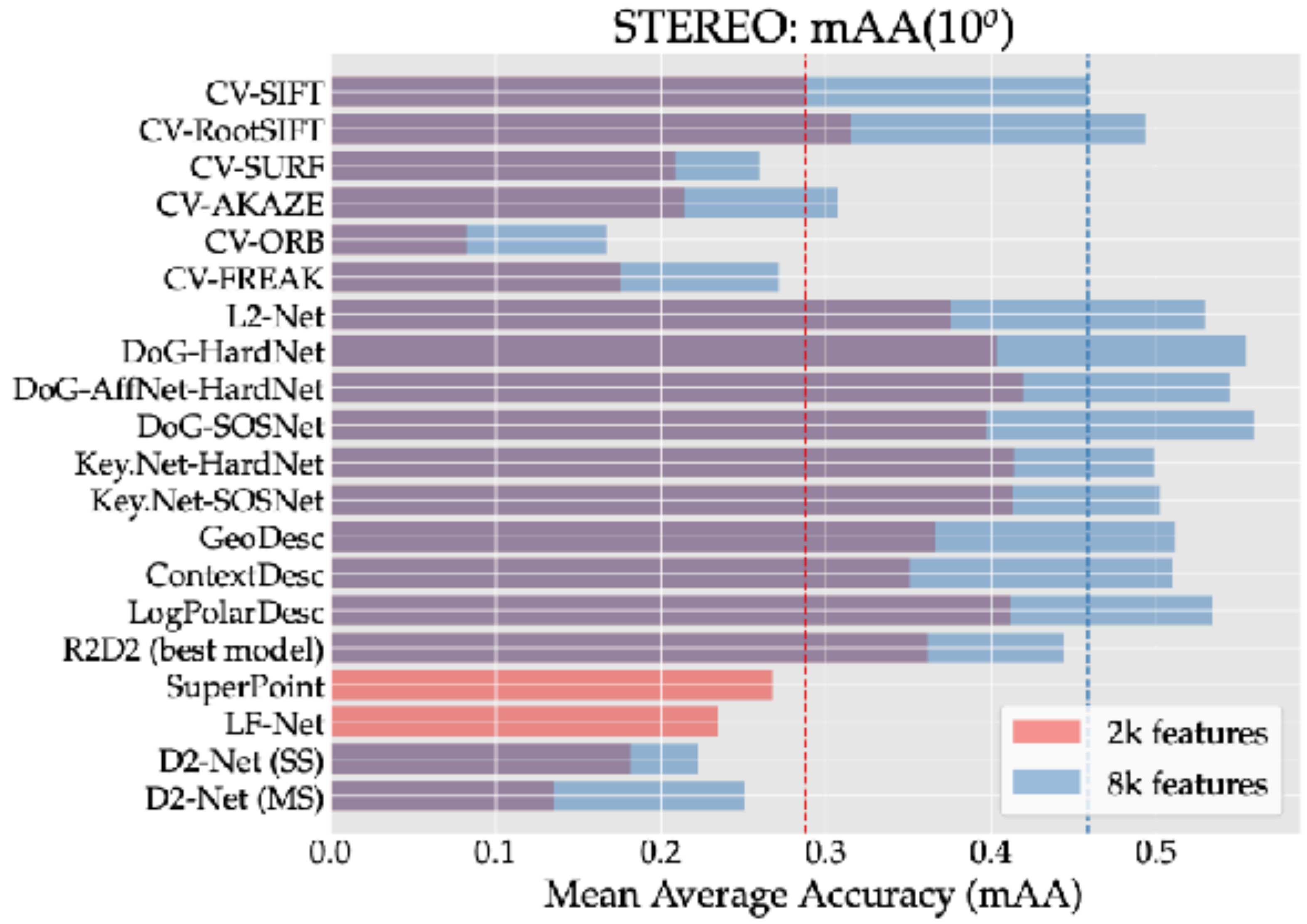


Fig. 17 Test – Stereo performance: 2k vs 8k features. We compare the results obtained with different methods using either 2k or 8k features – we use DEGENSAC, which performs better than other RANSAC variants under most circumstances. Dashed lines indicate SIFT's performance. For LF-Net and SuperPoint we do not include results with 8k features, as we failed to obtain meaningful results. For R2D2, we use the best model for each setting.

Checking results and methods

Start with baseline

- Find the method, which is the established "gold standard" and/or the baseline
- For image matching it is SIFT
- Check how it works
    - Note, that the "gold standard" might be complex and highly engineered
    - New methods might be simpler to implement
- Play selected (good and simple) methods
    - Maybe add their implementation to your ToDo

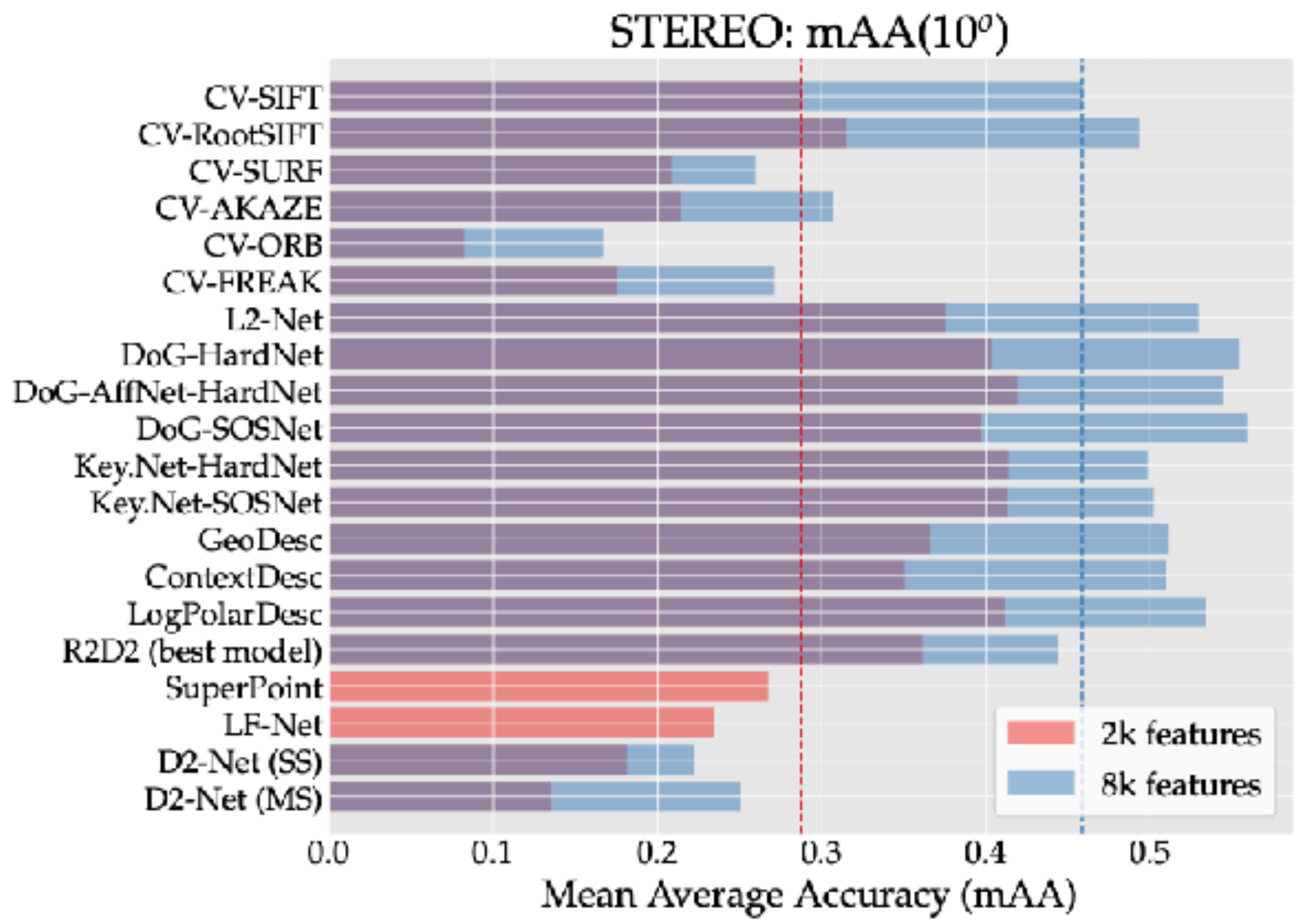


Fig. 17 **Test – Stereo performance: 2k vs 8k features.** We compare the results obtained with different methods using either 2k or 8k features – we use DEGENSAC, which performs better than other RANSAC variants under most circumstances. Dashed lines indicate SIFT's performance. For LF-Net and SuperPoint we do not include results with 8k features, as we failed to obtain meaningful results. For R2D2, we use the best model for each setting.

How to “play with method”?

How to “play with method”?

- First, try to make it work

How to “play with method”?

- First, try to make it work
  - Visualize, visualize, visualize

How to “play with method”?

- First, try to make it work
  - Visualize, visualize, visualize
- Once it works, try to break it. Noise, transforms, data, etc.

How to “play with method”?

- First, try to make it work
  - Visualize, visualize, visualize
- Once it works, try to break it. Noise, transforms, data, etc.
  - that it is how understand its limits.

How to “play with method”?

- First, try to make it work
  - Visualize, visualize, visualize
- Once it works, try to break it. Noise, transforms, data, etc.
  - that it is how understand its limits.
- Try to understand the ideas it is based on

How to “play with method”?

- First, try to make it work
  - Visualize, visualize, visualize
- Once it works, try to break it. Noise, transforms, data, etc.
  - that it is how understand its limits.
- Try to understand the ideas it is based on
- Only then try to implement it on your own

Playing with method

Case study

Playing with method
Case study

- I was working on some image-matching method with co-authors

Playing with method

Case study

- I was working on some image-matching method with co-authors
- They sent me the training code, I had to plug it into the benchmark

Playing with method
Case study

- I was working on some image-matching method with co-authors
- They sent me the training code, I had to plug it into the benchmark
- Instead I first wrote simple script to visualize keypoints

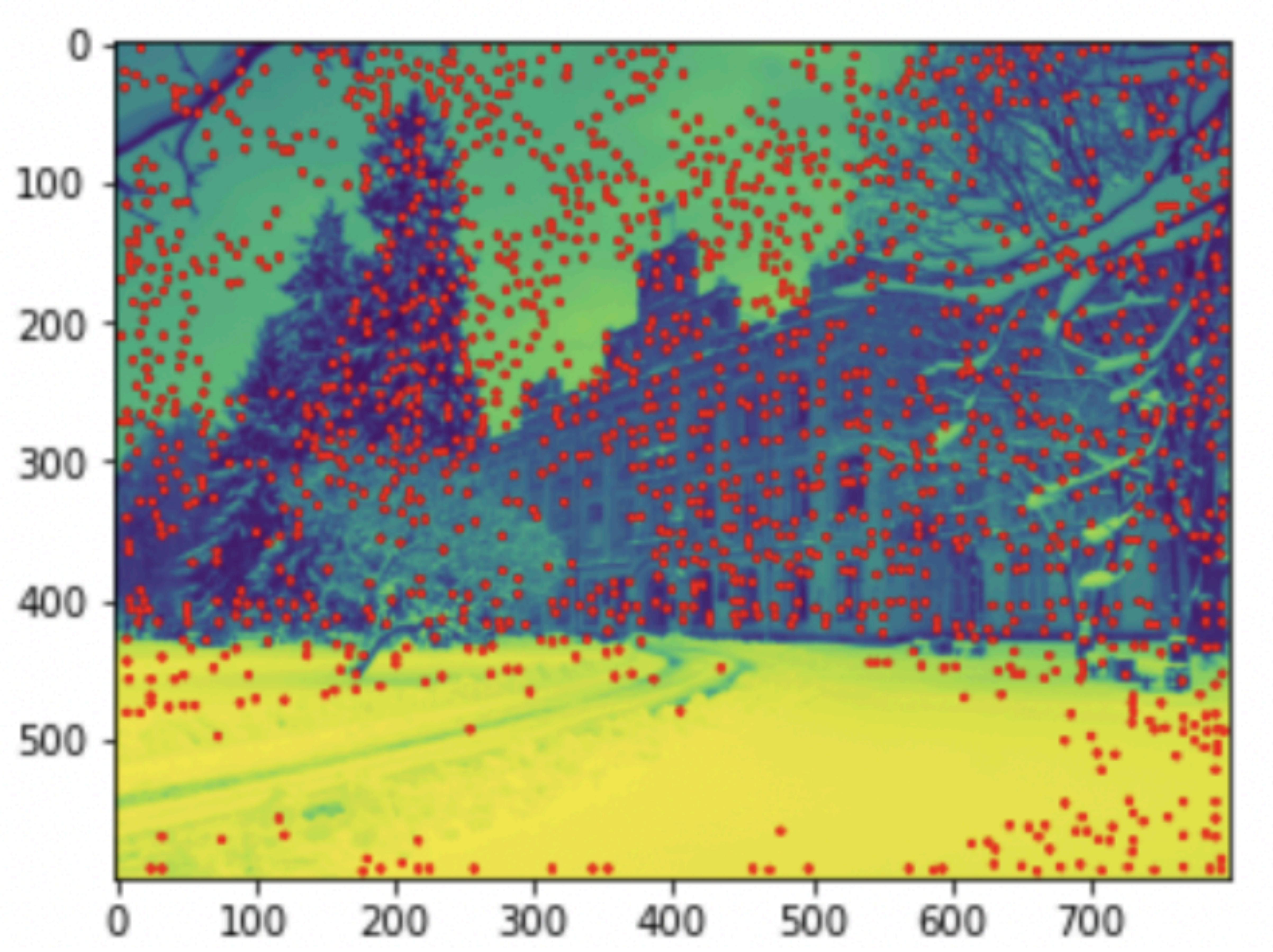
Playing with method

Case study

Playing with method

Case study

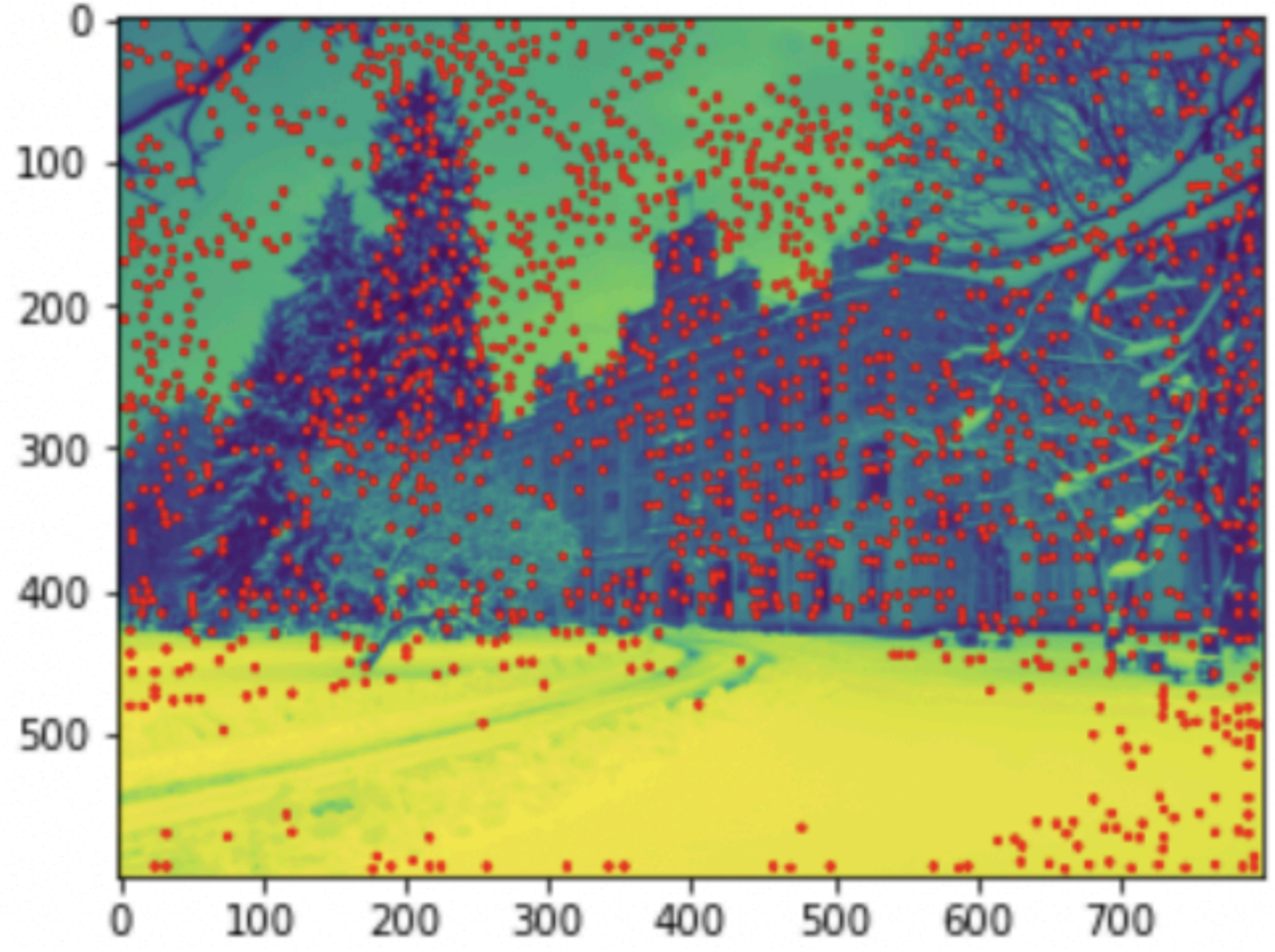
Model 1



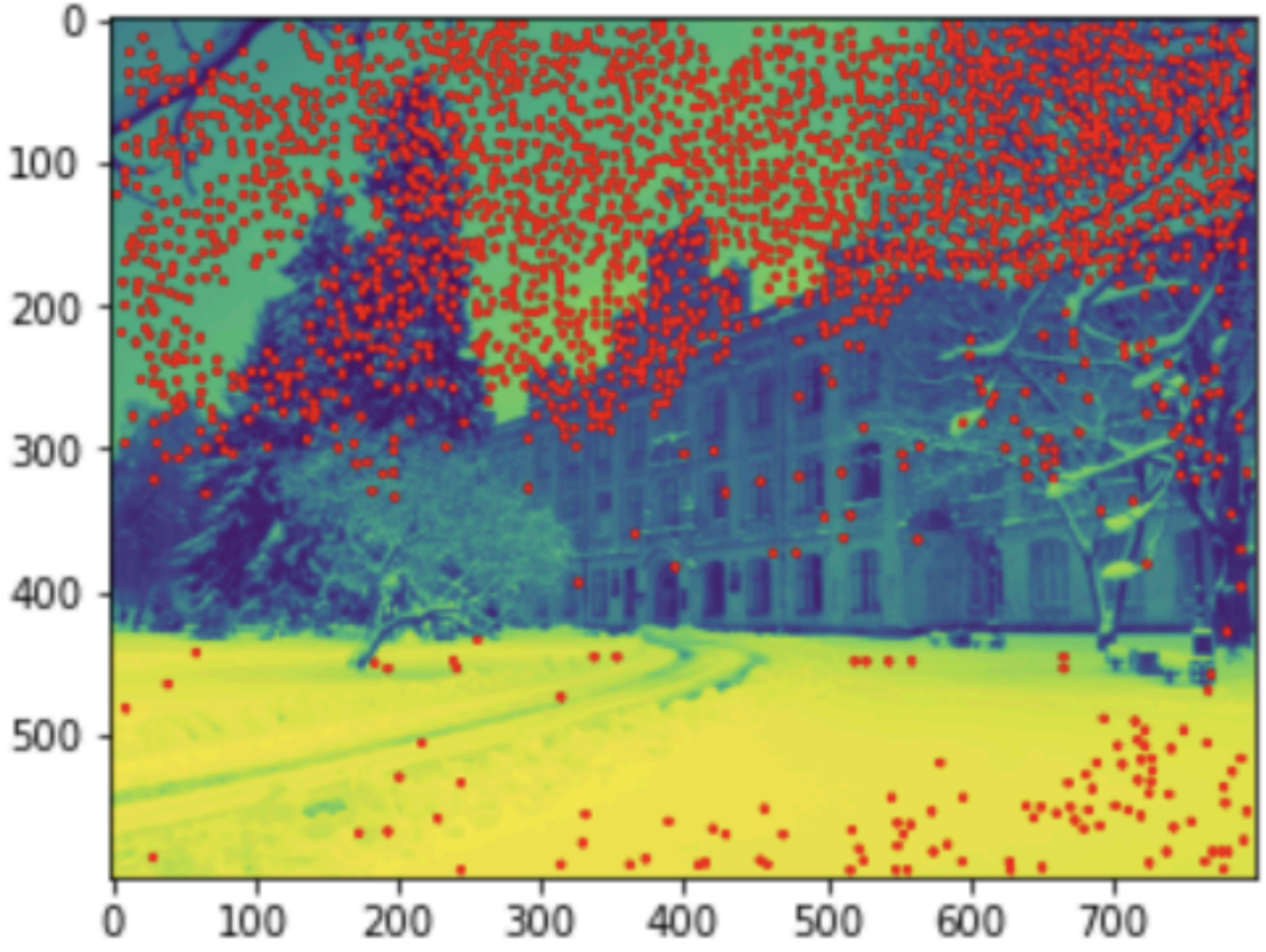
Playing with method

Case study

Model 1



Model 2

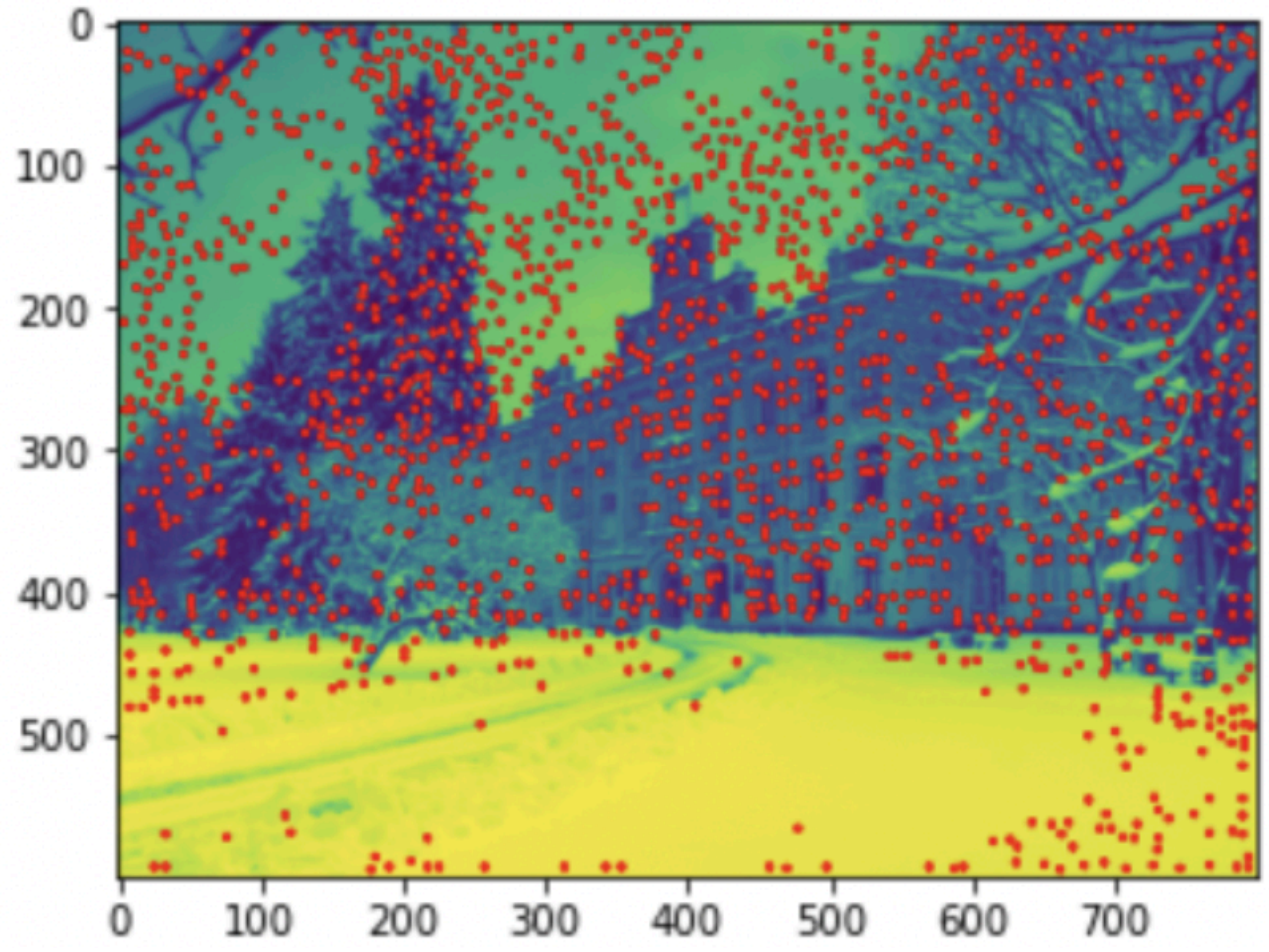


Playing with method

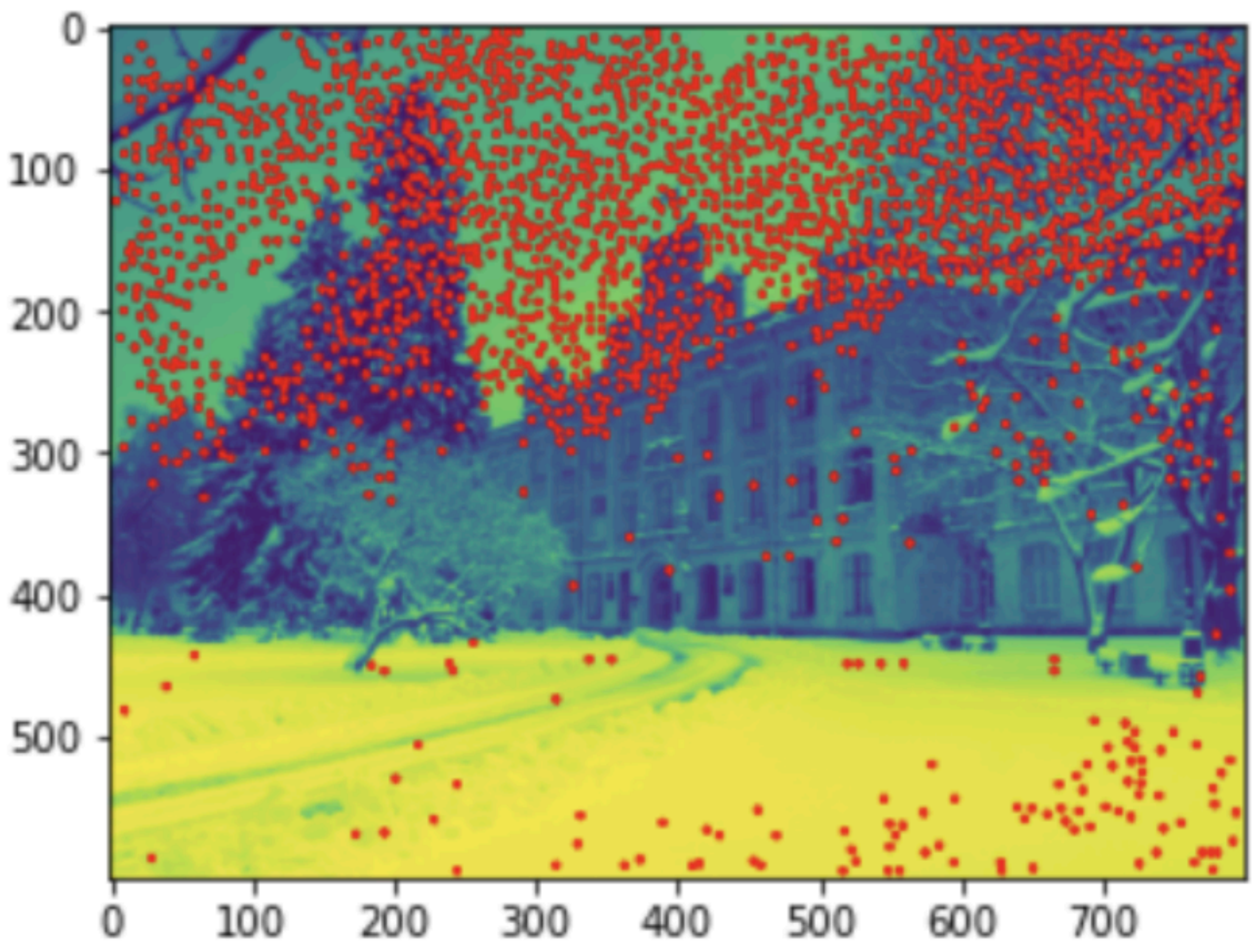
Case study

They were expected to look similar….

Model 1



Model 2

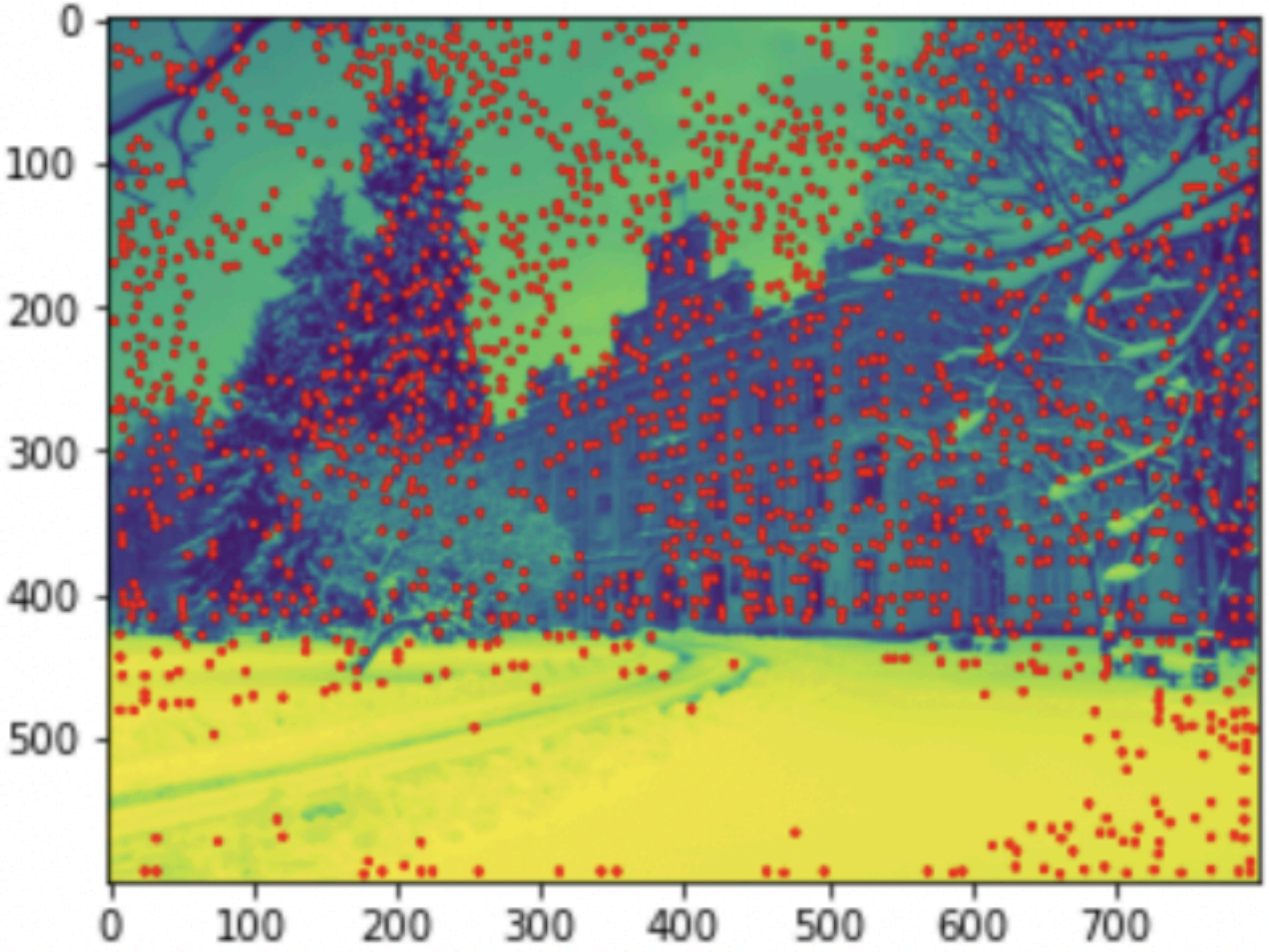


Playing with method

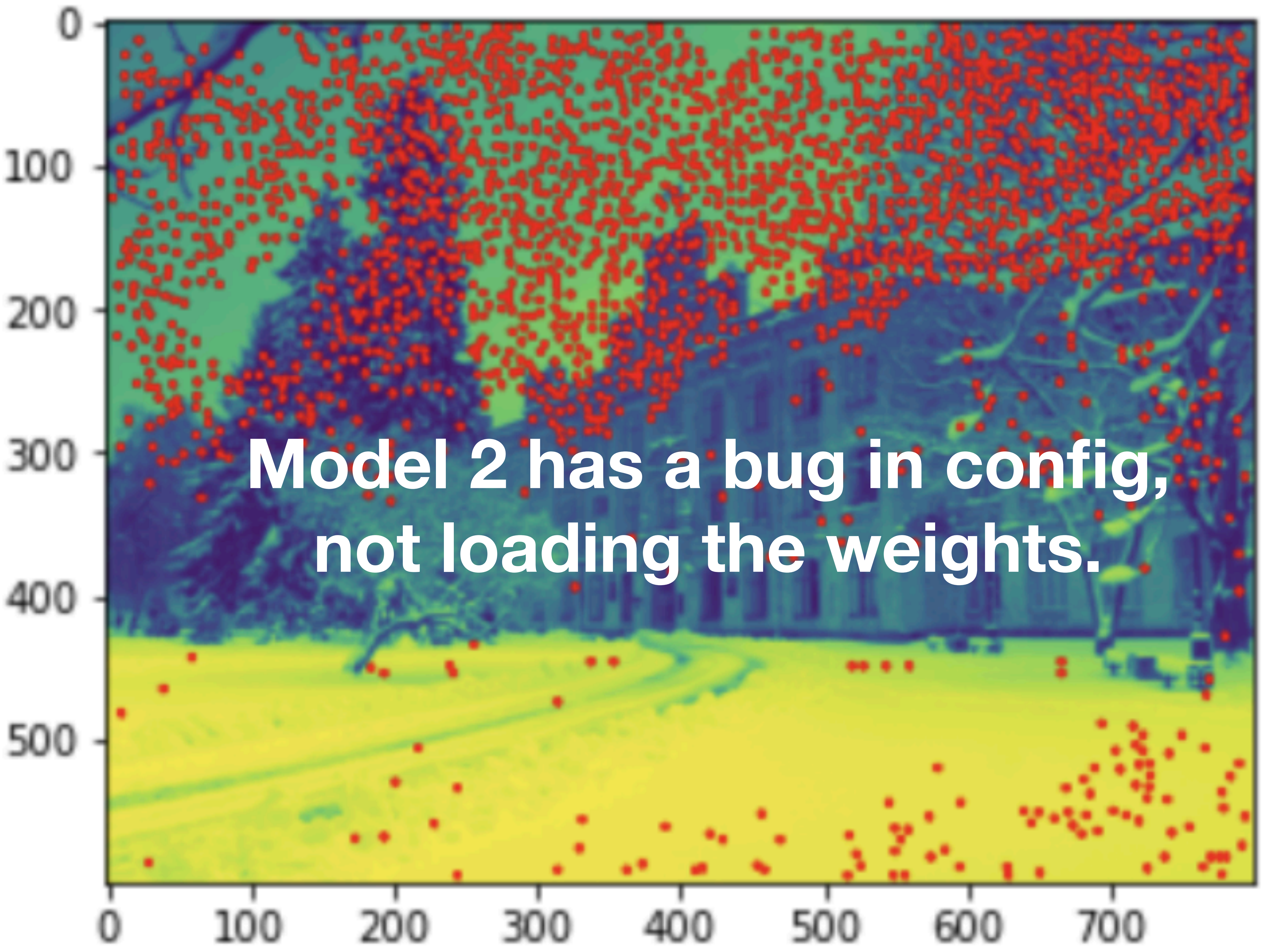
Case study

They were expected to look similar….

Model 1



Model 2



Playing with method

Case study

Playing with method

Case study

- If it is possible, gather some data yourself, not from benchmarks

Playing with method

Case study

- If it is possible, gather some data yourself, not from benchmarks
- Make it easy and fun to run the method on your data

Playing with method
Case study

- If it is possible, gather some data yourself, not from benchmarks
- Make it easy and fun to run the method on your data
- You can even build web demo for it

Playing with method

Case study

- If it is possible, gather some data yourself, not from benchmarks
- Make it easy and fun to run the method on your data
- You can even build web demo for it
- If it is not fun, it is not “playing”

Playing with method
Case study

- If it is possible, gather some data yourself, not from benchmarks
- Make it easy and fun to run the method on your data
- You can even build web demo for it
- If it is not fun, it is not “playing”
- You can think of this as “micro-benchmark”

Playing with method

Dmytro Mishkin
@ducha_aiki

R2D2 fails here, others are fine
9/

Перекласти твіт

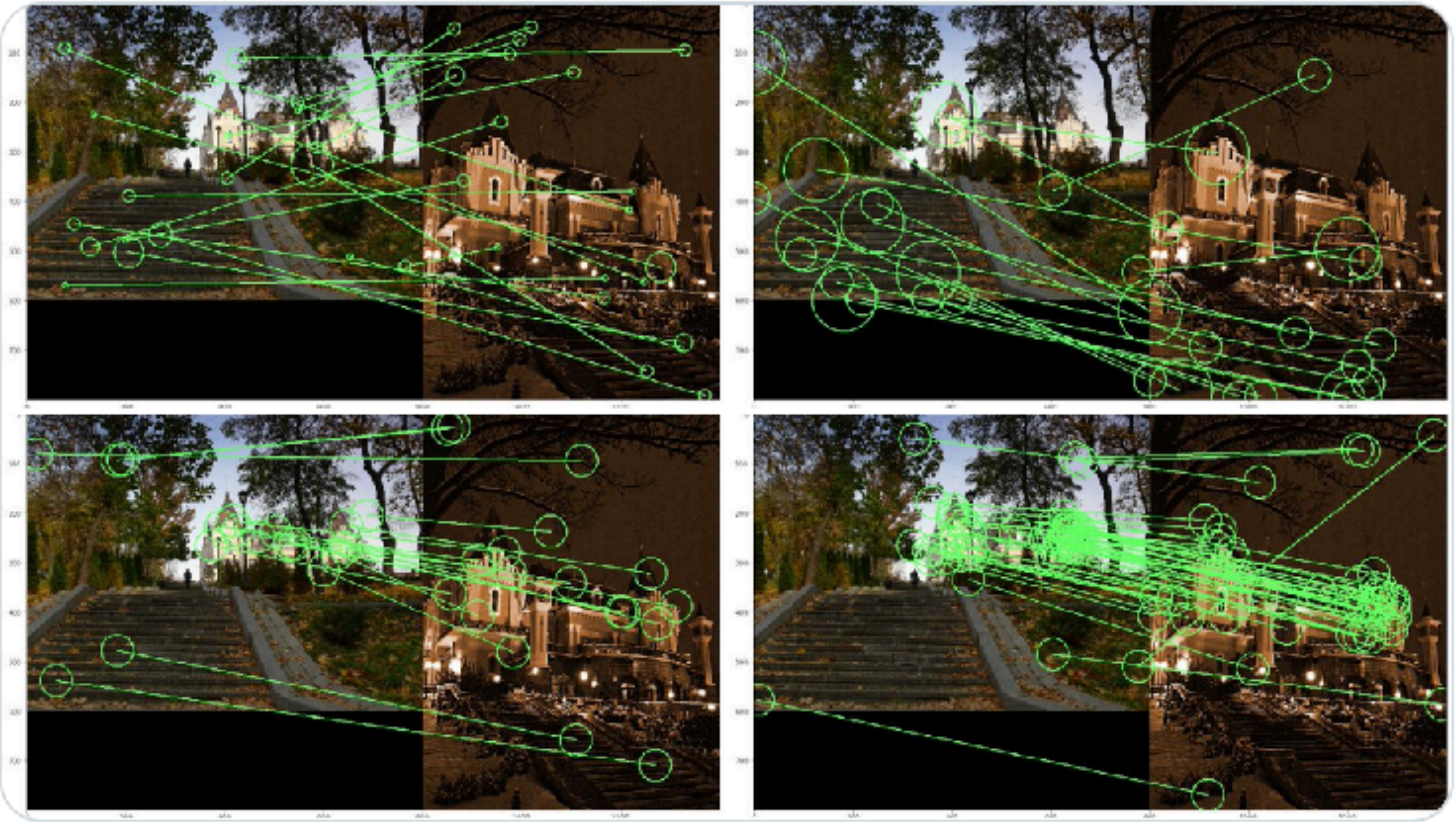
3:47 пп · 22 лист. 2020 · Twitter Web App

https://twitter.com/ducha_aiki/status/1330495426865344515

Dmytro Mishkin
@ducha_aiki

SIFT-HardNet and R2D2 fail here, SuperPoint rocks.
10/10
Overall, for random unknown upright pair of images with possible illumination change, I would go for SuperPoint, as least if we don't consider affine view synthesis.

Перекласти твіт



4:00 пп · 22 лист. 2020 · Twitter Web App

You know area well enough

and want to keep yourself up-to-date

You know area well enough

and want to keep yourself up-to-date

- You want to maximize useful information flow

You know area well enough

and want to keep yourself up-to-date

- You want to maximize useful information flow
- And minimize brain fuel & time spent

Where to look for papers on arXiv

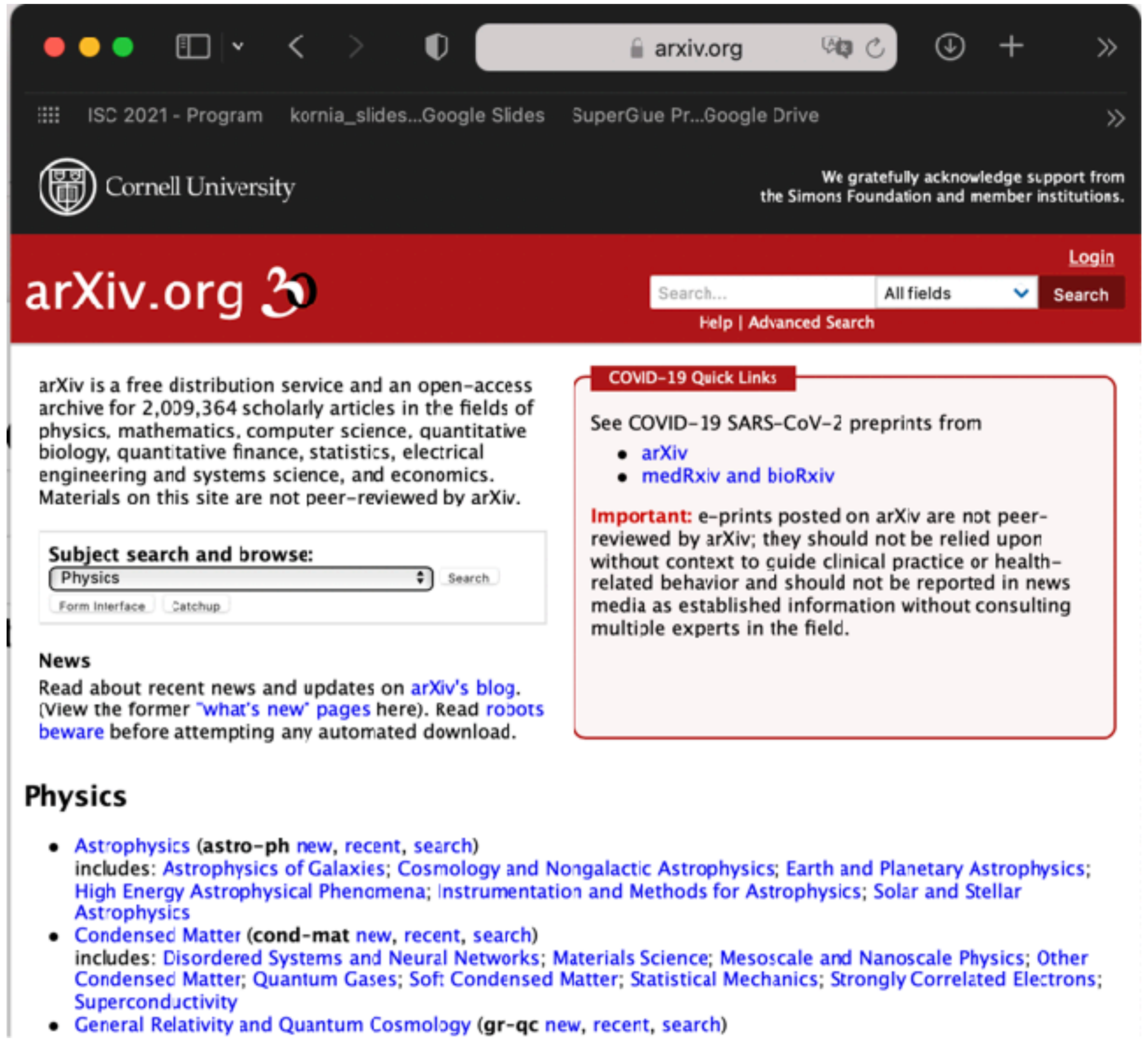
arXiv itself?

- arxiv.org

Where to look for papers on arXiv

arXiv itself?

- arxiv.org



Where to look for papers on arXiv

arXiv itself?

- arxiv.org

Computer Science

- Computing Research Repository (CoRR new, recent, search)
  includes (see detailed description): Artificial Intelligence; Computation and Language; Computational Complexity; Computational Engineering, Finance, and Science; Computational Geometry; Computer Science and Game Theory; Computer Vision and Pattern Recognition; Computers and Society; Cryptography and Security; Data Structures and Algorithms; Databases; Digital Libraries; Discrete Mathematics; Distributed, Parallel, and Cluster Computing; Emerging Technologies; Formal Languages and Automata Theory; General Literature; Graphics; Hardware Architecture; Human-Computer Interaction; Information Retrieval; Information Theory; Logic in Computer Science; Machine Learning; Mathematical Software; Multiagent Systems; Multimedia; Networking and Internet Architecture; Neural and Evolutionary Computing; Numerical Analysis; Operating Systems; Other Computer Science; Performance; Programming Languages; Robotics; Social and Information Networks; Software Engineering; Sound; Symbolic Computation; Systems and Control

Quantitative Biology

- Quantitative Biology (q-bio new, recent, search)
  includes (see detailed description): Biomolecules; Cell Behavior; Genomics; Molecular Networks; Neurons and Cognition; Other Quantitative Biology; Populations and Evolution; Quantitative Methods; Subcellular Processes; Tissues and Organs

Quantitative Finance

- Quantitative Finance (q-fin new, recent, search)
  includes (see detailed description): Computational Finance; Economics; General Finance; Mathematical Finance; Portfolio Management; Pricing of Securities; Risk Management; Statistical Finance; Trading and Market Microstructure

Statistics

- Statistics (stat new, recent, search)
  includes (see detailed description): Applications; Computation; Machine Learning; Methodology; Other Statistics; Statistics Theory

Where to look for papers on arXiv

arXiv itself?

- arxiv.org

Computer Science

- Computing Research Repository (CoRR new, recent, search)
  includes (see detailed description): Artificial Intelligence; Computation and Language; Computational Complexity; Computational Engineering, Finance, and Science; Computational Geometry; Computer Science and Game Theory; Computer Vision and Pattern Recognition; Computers and Society; Cryptography and Security; Data Structures and Algorithms; Databases; Digital Libraries; Discrete Mathematics; Distributed, Parallel, and Cluster Computing; Emerging Technologies; Formal Languages and Automata Theory; General Literature; Graphics; Hardware Architecture; Human-Computer Interaction; Information Retrieval; Information Theory; Logic in Computer Science; Machine Learning; Mathematical Software; Multiagent Systems; Multimedia; Networking and Internet Architecture; Neural and Evolutionary Computing; Numerical Analysis; Operating Systems; Other Computer Science; Performance; Programming Languages; Robotics; Social and Information Networks; Software Engineering; Sound; Symbolic Computation; Systems and Control

Quantitative Biology

- Quantitative Biology (q-bio new, recent, search)
  includes (see detailed description): Biomolecules; Cell Behavior; Genomics; Molecular Networks; Neurons and Cognition; Other Quantitative Biology; Populations and Evolution; Quantitative Methods; Subcellular Processes; Tissues and Organs

Quantitative Finance

- Quantitative Finance (q-fin new, recent, search)
  includes (see detailed description): Computational Finance; Economics; General Finance; Mathematical Finance; Portfolio Management; Pricing of Securities; Risk Management; Statistical Finance; Trading and Market Microstructure

Statistics

- Statistics (stat new, recent, search)
  includes (see detailed description): Applications; Computation; Machine Learning; Methodology; Other Statistics; Statistics Theory

- CS.Computer Vision and Pattern Recognition
- CS.Artificial Intelligence
- CS.Neural and Evolutionary Computing
- CS.Machine Learning
- CS.Robotics
- Stat.Machine Learning

What you can do on arXiv.org

arXiv.org > cs > cs.AI

Search... All fields Search

Help | Advanced Search

Artificial Intelligence

Authors and titles for recent submissions

- Tue, 25 Jan 2022
- Mon, 24 Jan 2022
- Fri, 21 Jan 2022
- Thu, 20 Jan 2022
- Wed, 19 Jan 2022

[total of 331 entries: **1-25** | 26-50 | 51-75 | 76-100 | ... | 326-331]
[showing 25 entries per page: fewer | more | all]

Tue, 25 Jan 2022 (showing first 25 of 79 entries)

[1] **arXiv:2201.09760** [**pdf**, **other**]

Multi-Graph Fusion Networks for Urban Region Embedding
Shangbin Wu, Xu Yan, Xiaoliang Fan, Shirui Pan, Shichao Zhu, Chuanpan Zheng, Ming Cheng, Cheng Wang
Subjects: **Artificial Intelligence (cs.AI)**

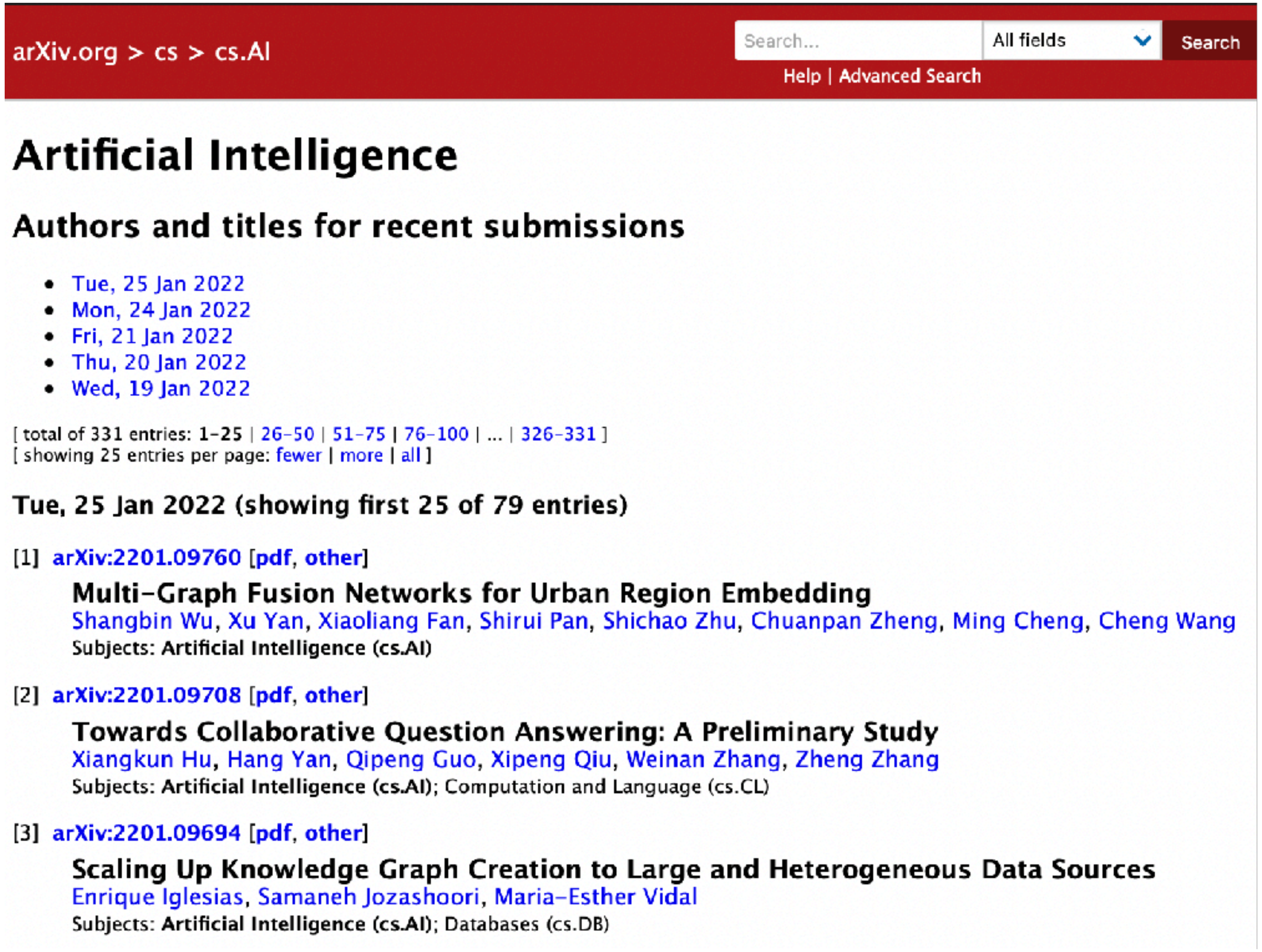
[2] **arXiv:2201.09708** [**pdf**, **other**]

Towards Collaborative Question Answering: A Preliminary Study
Xiangkun Hu, Hang Yan, Qipeng Guo, Xipeng Qiu, Weinan Zhang, Zheng Zhang
Subjects: **Artificial Intelligence (cs.AI)**; Computation and Language (cs.CL)

[3] **arXiv:2201.09694** [**pdf**, **other**]

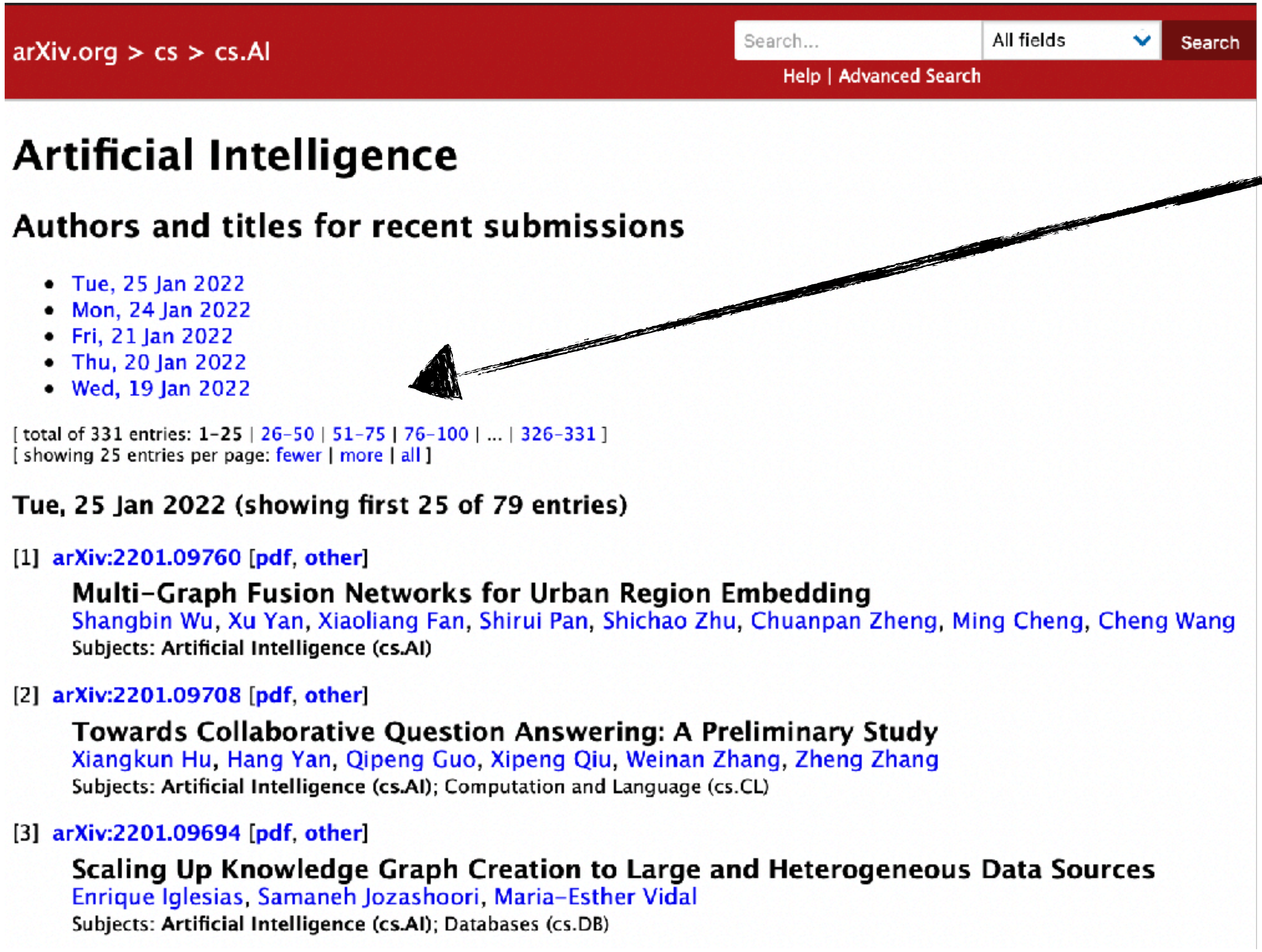
Scaling Up Knowledge Graph Creation to Large and Heterogeneous Data Sources
Enrique Iglesias, Samaneh Jozashoori, Maria-Esther Vidal
Subjects: **Artificial Intelligence (cs.AI)**; Databases (cs.DB)

What you can do on arXiv.org



Dates, number of entries per page

What you can do on arXiv.org


arXiv.org > cs > cs.AI

Search... | All fields | Search

Help | Advanced Search

Artificial Intelligence

Authors and titles for recent submissions

- Tue, 25 Jan 2022
- Mon, 24 Jan 2022
- Fri, 21 Jan 2022
- Thu, 20 Jan 2022
- Wed, 19 Jan 2022

[total of 331 entries: 1-25 | 26-50 | 51-75 | 76-100 | ... | 326-331]
[showing 25 entries per page: fewer | more | all]

Tue, 25 Jan 2022 (showing first 25 of 79 entries)

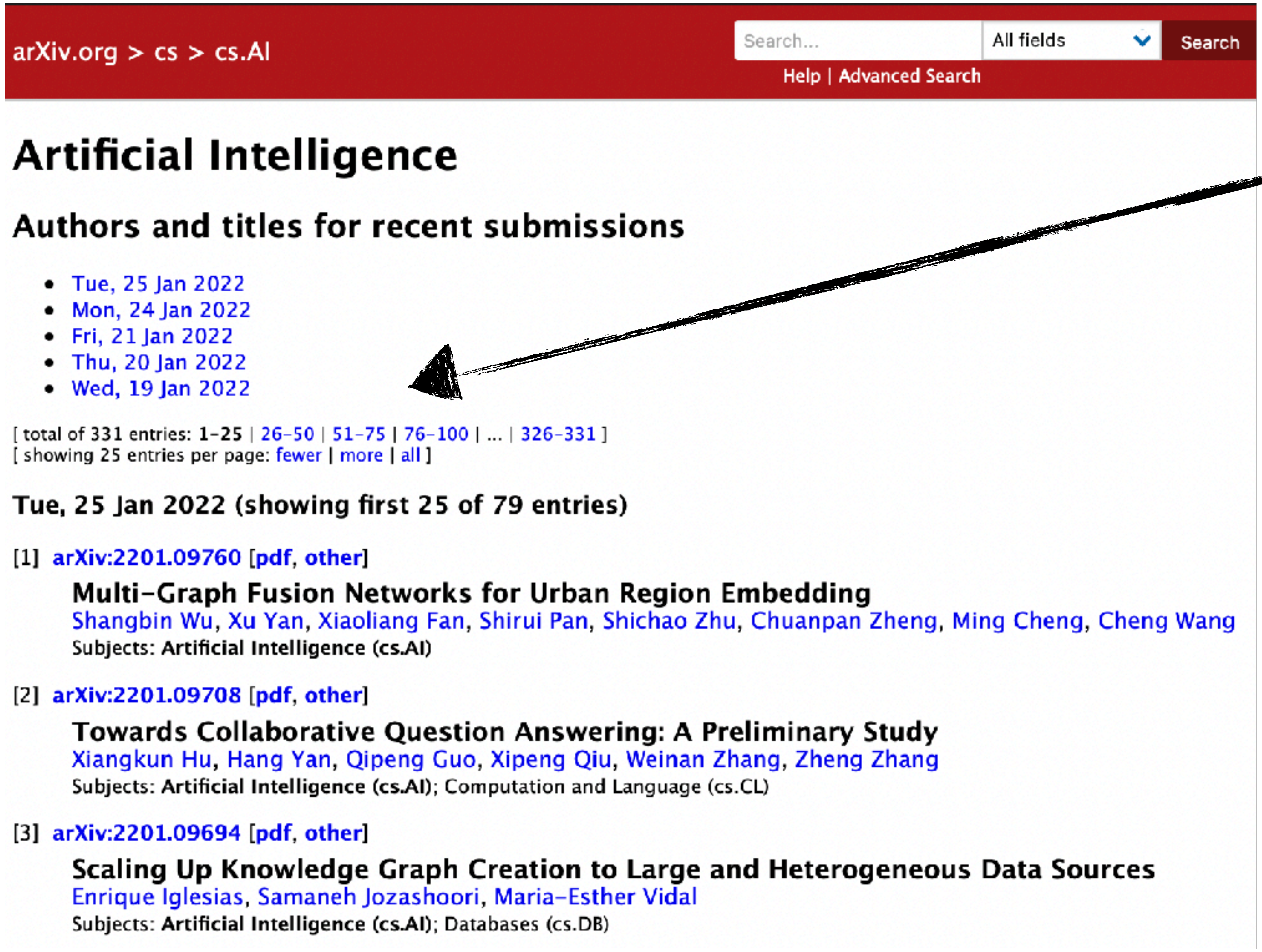
[1] arXiv:2201.09760 [pdf, other]
Multi-Graph Fusion Networks for Urban Region Embedding
Shangbin Wu, Xu Yan, Xiaoliang Fan, Shirui Pan, Shichao Zhu, Chuanpan Zheng, Ming Cheng, Cheng Wang
Subjects: Artificial Intelligence (cs.AI)

[2] arXiv:2201.09708 [pdf, other]
Towards Collaborative Question Answering: A Preliminary Study
Xiangkun Hu, Hang Yan, Qipeng Guo, Xipeng Qiu, Weinan Zhang, Zheng Zhang
Subjects: Artificial Intelligence (cs.AI); Computation and Language (cs.CL)

[3] arXiv:2201.09694 [pdf, other]
Scaling Up Knowledge Graph Creation to Large and Heterogeneous Data Sources
Enrique Iglesias, Samaneh Jozashoori, Maria-Esther Vidal
Subjects: Artificial Intelligence (cs.AI); Databases (cs.DB)

Dates, number of entries per page

What you can do on arXiv.org


arXiv.org > cs > cs.AI

Search... All fields Search

Help | Advanced Search

Artificial Intelligence

Authors and titles for recent submissions

- Tue, 25 Jan 2022
- Mon, 24 Jan 2022
- Fri, 21 Jan 2022
- Thu, 20 Jan 2022
- Wed, 19 Jan 2022

[total of 331 entries: 1-25 | 26-50 | 51-75 | 76-100 | ... | 326-331]
[showing 25 entries per page: fewer | more | all]

Tue, 25 Jan 2022 (showing first 25 of 79 entries)

[1] arXiv:2201.09760 [pdf, other]

Multi-Graph Fusion Networks for Urban Region Embedding
Shangbin Wu, Xu Yan, Xiaoliang Fan, Shirui Pan, Shichao Zhu, Chuanpan Zheng, Ming Cheng, Cheng Wang
Subjects: **Artificial Intelligence (cs.AI)**

[2] arXiv:2201.09708 [pdf, other]

Towards Collaborative Question Answering: A Preliminary Study
Xiangkun Hu, Hang Yan, Qipeng Guo, Xipeng Qiu, Weinan Zhang, Zheng Zhang
Subjects: **Artificial Intelligence (cs.AI)**; Computation and Language (cs.CL)

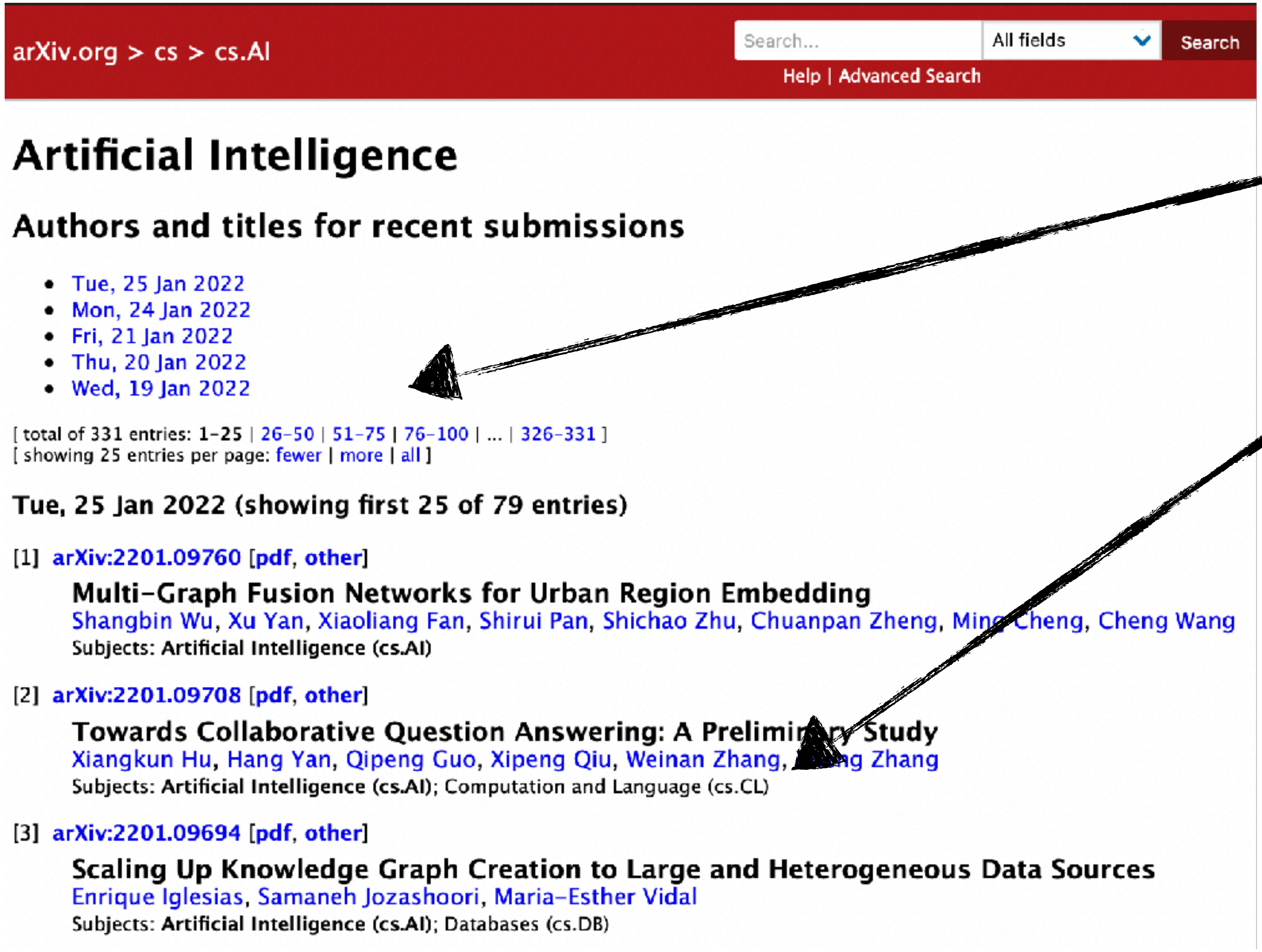
[3] arXiv:2201.09694 [pdf, other]

Scaling Up Knowledge Graph Creation to Large and Heterogeneous Data Sources
Enrique Iglesias, Samaneh Jozashoori, Maria-Esther Vidal
Subjects: **Artificial Intelligence (cs.AI)**; Databases (cs.DB)

Dates, number of entries per page

Some papers appear
on several lists

What you can do on arXiv.org


arXiv.org > cs > cs.AI

Search... All fields Search

Help | Advanced Search

Artificial Intelligence

Authors and titles for recent submissions

- Tue, 25 Jan 2022
- Mon, 24 Jan 2022
- Fri, 21 Jan 2022
- Thu, 20 Jan 2022
- Wed, 19 Jan 2022

[total of 331 entries: **1-25** | 26-50 | 51-75 | 76-100 | ... | 326-331]
[showing 25 entries per page: fewer | more | all]

Tue, 25 Jan 2022 (showing first 25 of 79 entries)

[1] **arXiv:2201.09760** [**pdf**, **other**]

Multi-Graph Fusion Networks for Urban Region Embedding
Shangbin Wu, Xu Yan, Xiaoliang Fan, Shirui Pan, Shichao Zhu, Chuanpan Zheng, Ming Cheng, Cheng Wang
Subjects: **Artificial Intelligence (cs.AI)**

[2] **arXiv:2201.09708** [**pdf**, **other**]

Towards Collaborative Question Answering: A Preliminary Study
Xiangkun Hu, Hang Yan, Qipeng Guo, Xipeng Qiu, Weinan Zhang, ng Zhang
Subjects: **Artificial Intelligence (cs.AI)**; Computation and Language (cs.CL)

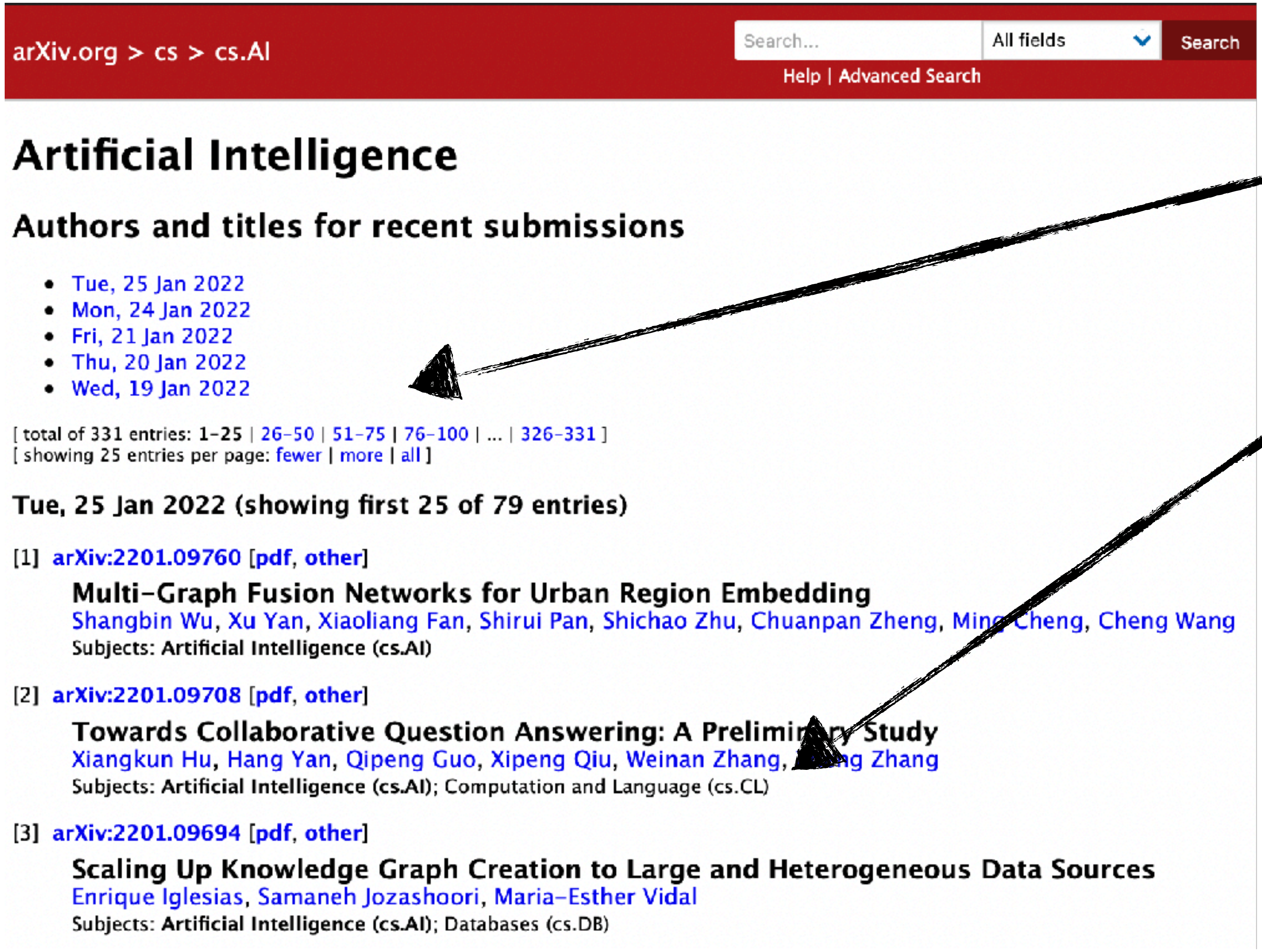
[3] **arXiv:2201.09694** [**pdf**, **other**]

Scaling Up Knowledge Graph Creation to Large and Heterogeneous Data Sources
Enrique Iglesias, Samaneh Jozashoori, Maria-Esther Vidal
Subjects: **Artificial Intelligence (cs.AI)**; Databases (cs.DB)

Dates, number of entries per page

Some papers appear
on several lists

What you can do on arXiv.org



arXiv.org > cs > cs.AI

Search... | All fields | Search

Help | Advanced Search

Artificial Intelligence

Authors and titles for recent submissions

- Tue, 25 Jan 2022
- Mon, 24 Jan 2022
- Fri, 21 Jan 2022
- Thu, 20 Jan 2022
- Wed, 19 Jan 2022

[total of 331 entries: 1-25 | 26-50 | 51-75 | 76-100 | ... | 326-331]
[showing 25 entries per page: fewer | more | all]

Tue, 25 Jan 2022 (showing first 25 of 79 entries)

[1] arXiv:2201.09760 [pdf, other]
Multi-Graph Fusion Networks for Urban Region Embedding
Shangbin Wu, Xu Yan, Xiaoliang Fan, Shirui Pan, Shichao Zhu, Chuanpan Zheng, Ming Cheng, Cheng Wang
Subjects: **Artificial Intelligence (cs.AI)**

[2] arXiv:2201.09708 [pdf, other]
Towards Collaborative Question Answering: A Preliminary Study
Xiangkun Hu, Hang Yan, Qipeng Guo, Xipeng Qiu, Weinan Zhang, [illegible]ng Zhang
Subjects: **Artificial Intelligence (cs.AI)**; Computation and Language (cs.CL)

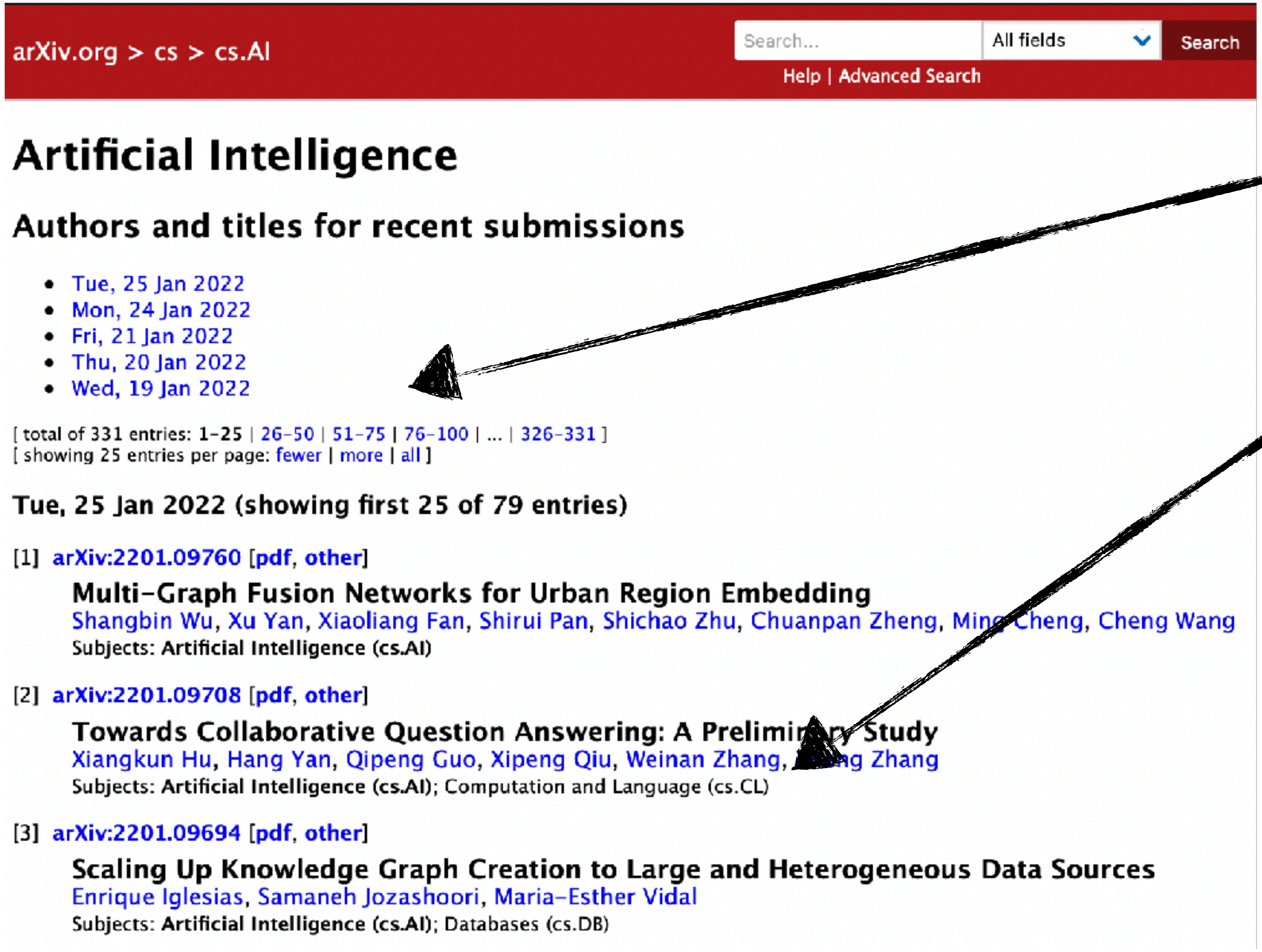
[3] arXiv:2201.09694 [pdf, other]
Scaling Up Knowledge Graph Creation to Large and Heterogeneous Data Sources
Enrique Iglesias, Samaneh Jozashoori, Maria-Esther Vidal
Subjects: **Artificial Intelligence (cs.AI)**; Databases (cs.DB)

Dates, number of entries per page

Some papers appear
on several lists

You can see only title and
authors

What you can do on arXiv.org



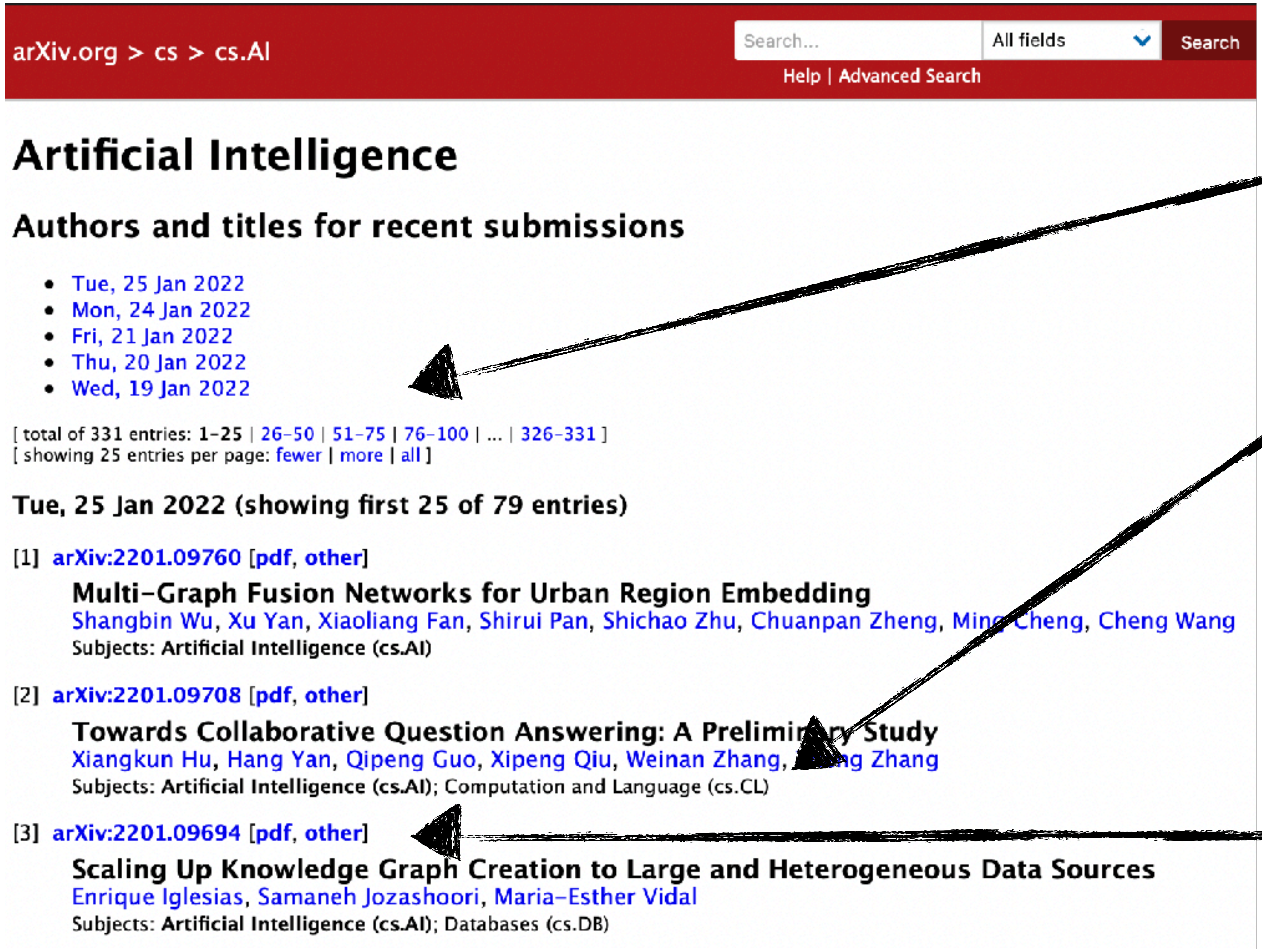
Dates, number of entries per page

Some papers appear on several lists

You can see only title and authors

“other” means latex source, useful for copying tables, images, etc

What you can do on arXiv.org



Dates, number of entries per page

Some papers appear on several lists

You can see only title and authors

“other” means latex source, useful for copying tables, images, etc

What you can do on arXiv.org

“other" format

arXiv.org

Search...

All fields

Search

Help | Advanced Search

Format selector for arXiv:2201.09760v1

Set cookies: if your browser supports cookies you can configure your default format.

PDF

Now includes fonts, see our PDF help. Download PDF

Source

Delivered as a **gzipped tar** (.tar.gz) file if there are multiple files, otherwise as a **PDF** file, or a **gzipped TeX, DVI, PostScript or HTML** (.gz, .dvi.gz, .ps.gz or .html.gz) file depending on submission format. [Download source]

For more information consult the viewing help.

Note: Many of the formats above are served gzipped (`Content-Encoding: x-gzip`). Your browser may silently uncompress after downloading so the files you see saved may appear uncompressed.

Please report any problems to arXiv admins, being certain to state the paper identifier, explicit options you've selected and that you're using arxiv.org.

What you can do on arXiv.org

“other" format

arxiv-paper — -zsh — 80×24

```
Last login: Tue Jan 25 15:39:25 on ttys000
arm64
(python39) oldufo@Dmytros-MacBook-Air ~ % cd Downloads/arxiv-paper
(python39) oldufo@Dmytros-MacBook-Air arxiv-paper % ls
2201.09760
(python39) oldufo@Dmytros-MacBook-Air arxiv-paper % mv 2201.09760 2201.09760.tgz

(python39) oldufo@Dmytros-MacBook-Air arxiv-paper % tar -xzf 2201.09760.tgz
(python39) oldufo@Dmytros-MacBook-Air arxiv-paper % ls
0_abstract.tex          4_Evaluation.tex         ijcai22.bib
0_authors.tex           5_RelatedWork.tex        ijcai22.sty
1_Introduction.tex      6_Conclusion.tex         ijcai22.tex
2201.09760.tgz          Figure                   named.bst
2_ProblemStatement.tex  ijcai22-multiauthor.pdf
3_Methodology.tex       ijcai22.bbl
(python39) oldufo@Dmytros-MacBook-Air arxiv-paper %
```

What you can do on arXiv.org

“other" format

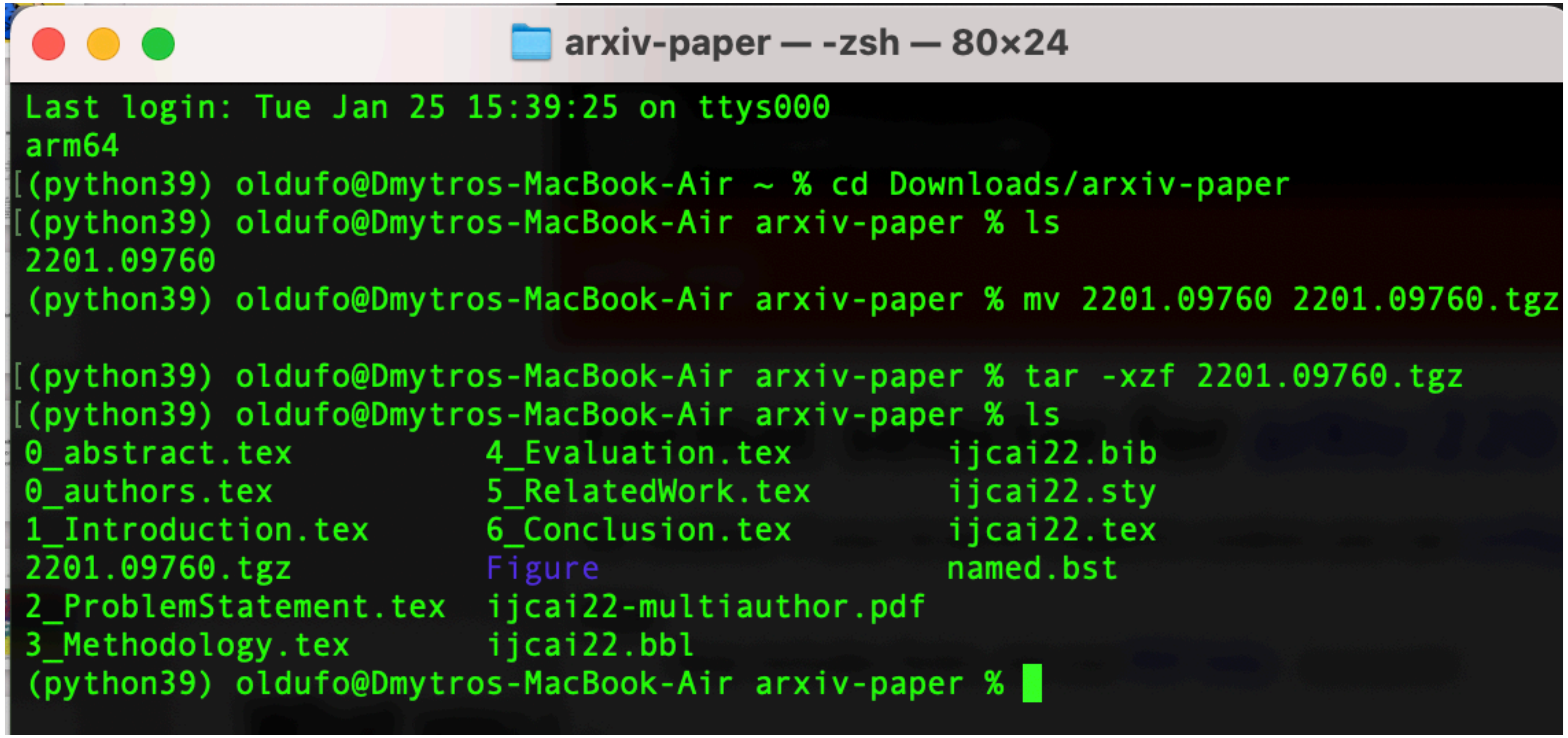
```
arxiv-paper — -zsh — 80×24
Last login: Tue Jan 25 15:39:25 on ttys000
arm64
(python39) oldufo@Dmytros-MacBook-Air ~ % cd Downloads/arxiv-paper
(python39) oldufo@Dmytros-MacBook-Air arxiv-paper % ls
2201.09760
(python39) oldufo@Dmytros-MacBook-Air arxiv-paper % mv 2201.09760 2201.09760.tgz

(python39) oldufo@Dmytros-MacBook-Air arxiv-paper % tar -xzf 2201.09760.tgz
(python39) oldufo@Dmytros-MacBook-Air arxiv-paper % ls
0_abstract.tex          4_Evaluation.tex        ijcai22.bib
0_authors.tex           5_RelatedWork.tex       ijcai22.sty
1_Introduction.tex      6_Conclusion.tex        ijcai22.tex
2201.09760.tgz          Figure                  named.bst
2_ProblemStatement.tex  ijcai22-multiauthor.pdf
3_Methodology.tex       ijcai22.bbl
(python39) oldufo@Dmytros-MacBook-Air arxiv-paper %
```

- arXiv downloads “other" as a file without extension

What you can do on arXiv.org

“other" format



```
arxiv-paper — -zsh — 80×24
Last login: Tue Jan 25 15:39:25 on ttys000
arm64
(python39) oldufo@Dmytros-MacBook-Air ~ % cd Downloads/arxiv-paper
(python39) oldufo@Dmytros-MacBook-Air arxiv-paper % ls
2201.09760
(python39) oldufo@Dmytros-MacBook-Air arxiv-paper % mv 2201.09760 2201.09760.tgz

(python39) oldufo@Dmytros-MacBook-Air arxiv-paper % tar -xzf 2201.09760.tgz
(python39) oldufo@Dmytros-MacBook-Air arxiv-paper % ls
0_abstract.tex          4_Evaluation.tex        ijcai22.bib
0_authors.tex           5_RelatedWork.tex       ijcai22.sty
1_Introduction.tex      6_Conclusion.tex        ijcai22.tex
2201.09760.tgz          Figure                  named.bst
2_ProblemStatement.tex  ijcai22-multiauthor.pdf
3_Methodology.tex       ijcai22.bbl
(python39) oldufo@Dmytros-MacBook-Air arxiv-paper %
```

- arXiv downloads “other" as a file without extension
- it is actually tar.gz file

What you can do on arXiv.org

“other" format

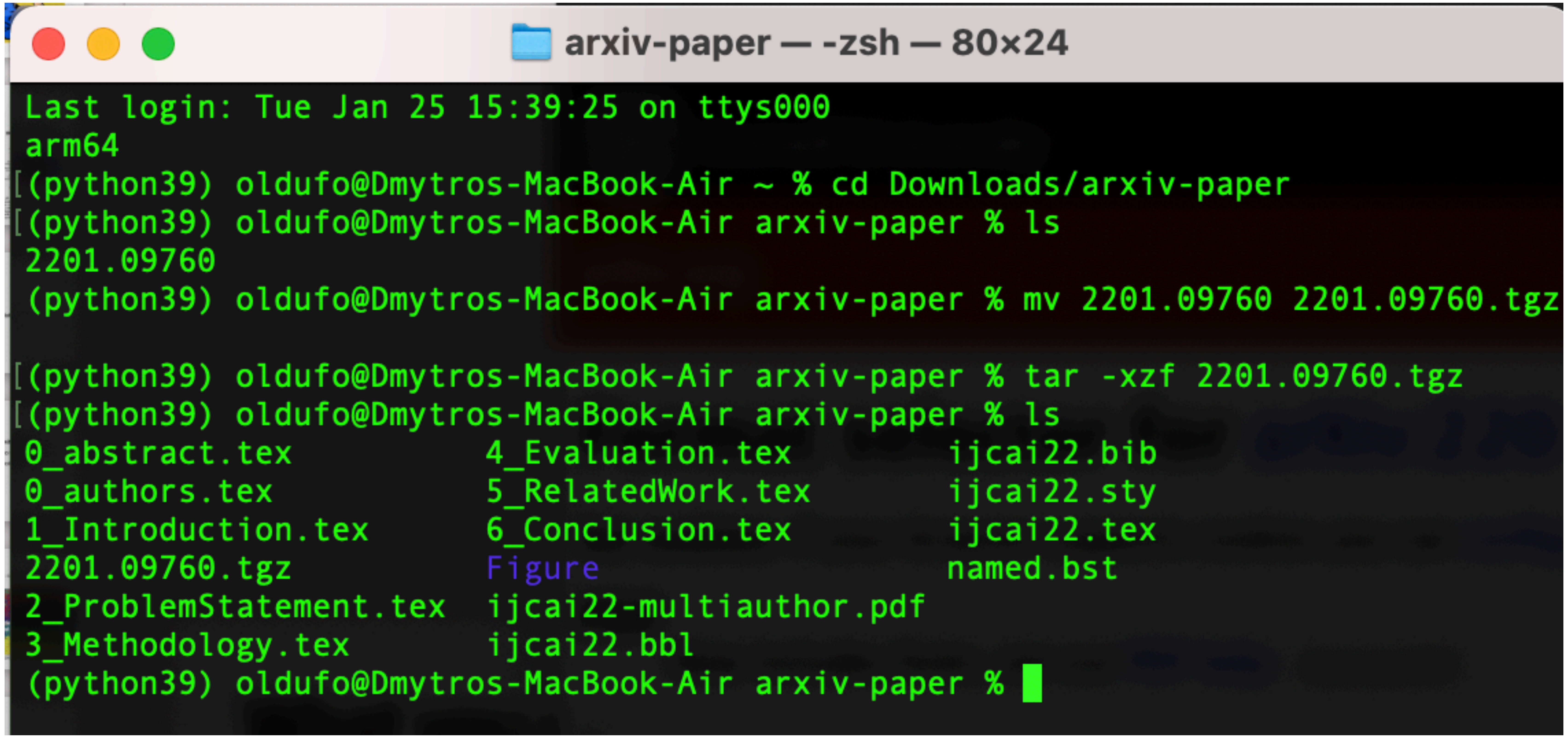
```
arxiv-paper — -zsh — 80×24
Last login: Tue Jan 25 15:39:25 on ttys000
arm64
(python39) oldufo@Dmytros-MacBook-Air ~ % cd Downloads/arxiv-paper
(python39) oldufo@Dmytros-MacBook-Air arxiv-paper % ls
2201.09760
(python39) oldufo@Dmytros-MacBook-Air arxiv-paper % mv 2201.09760 2201.09760.tgz

(python39) oldufo@Dmytros-MacBook-Air arxiv-paper % tar -xzf 2201.09760.tgz
(python39) oldufo@Dmytros-MacBook-Air arxiv-paper % ls
0_abstract.tex          4_Evaluation.tex          ijcai22.bib
0_authors.tex           5_RelatedWork.tex         ijcai22.sty
1_Introduction.tex      6_Conclusion.tex          ijcai22.tex
2201.09760.tgz          Figure                    named.bst
2_ProblemStatement.tex  ijcai22-multiauthor.pdf
3_Methodology.tex       ijcai22.bbl
(python39) oldufo@Dmytros-MacBook-Air arxiv-paper %
```

- arXiv downloads “other" as a file without extension
- it is actually tar.gz file
- It will unpack to the current dir, so better to create a new dir

What you can do on arXiv.org

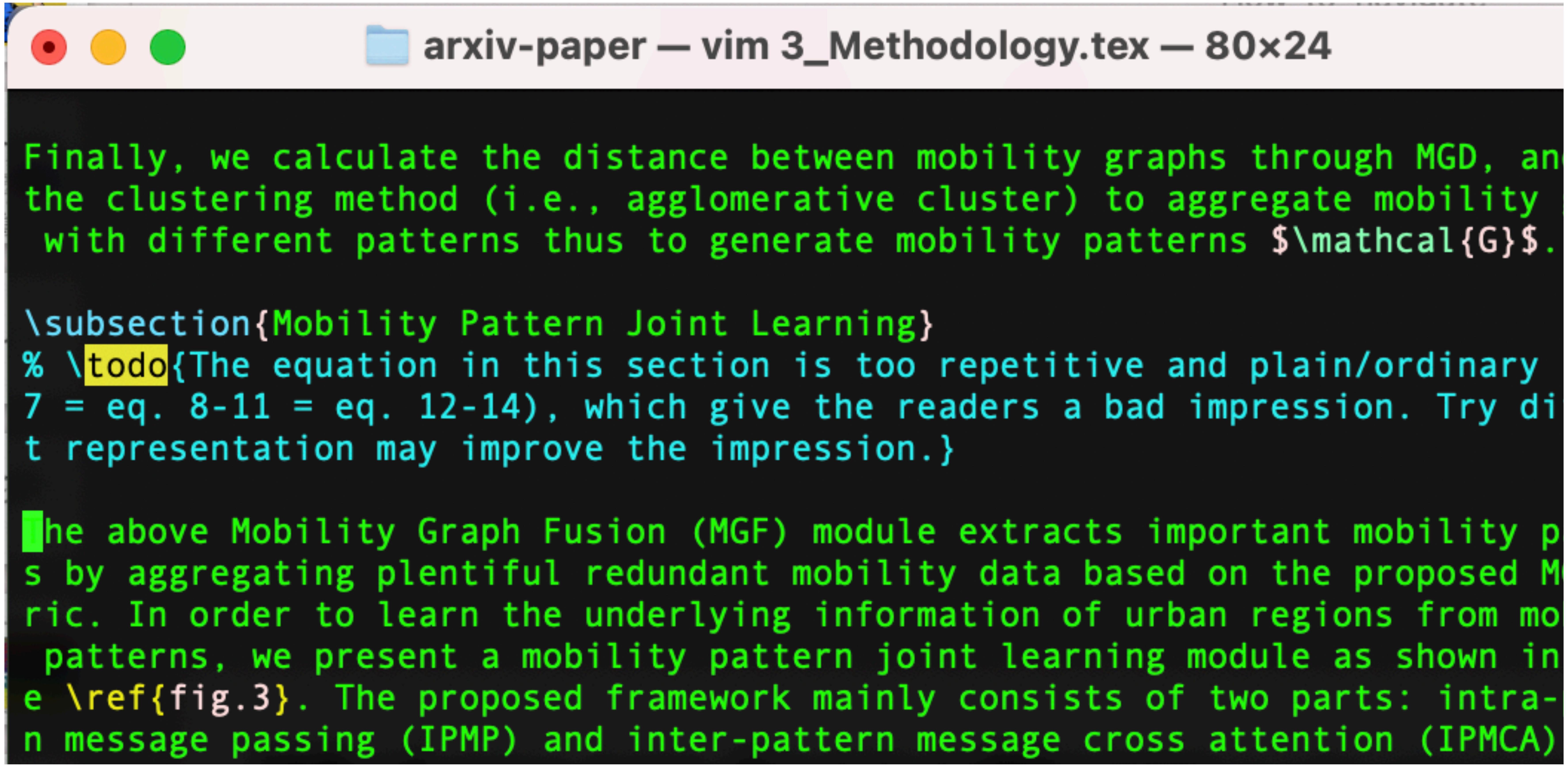
“other" format



- arXiv downloads “other" as a file without extension
- it is actually tar.gz file
- It will unpack to the current dir, so better to create a new dir
- You can find interesting points and todos there, if authors haven’t deleted them

What you can do on arXiv.org

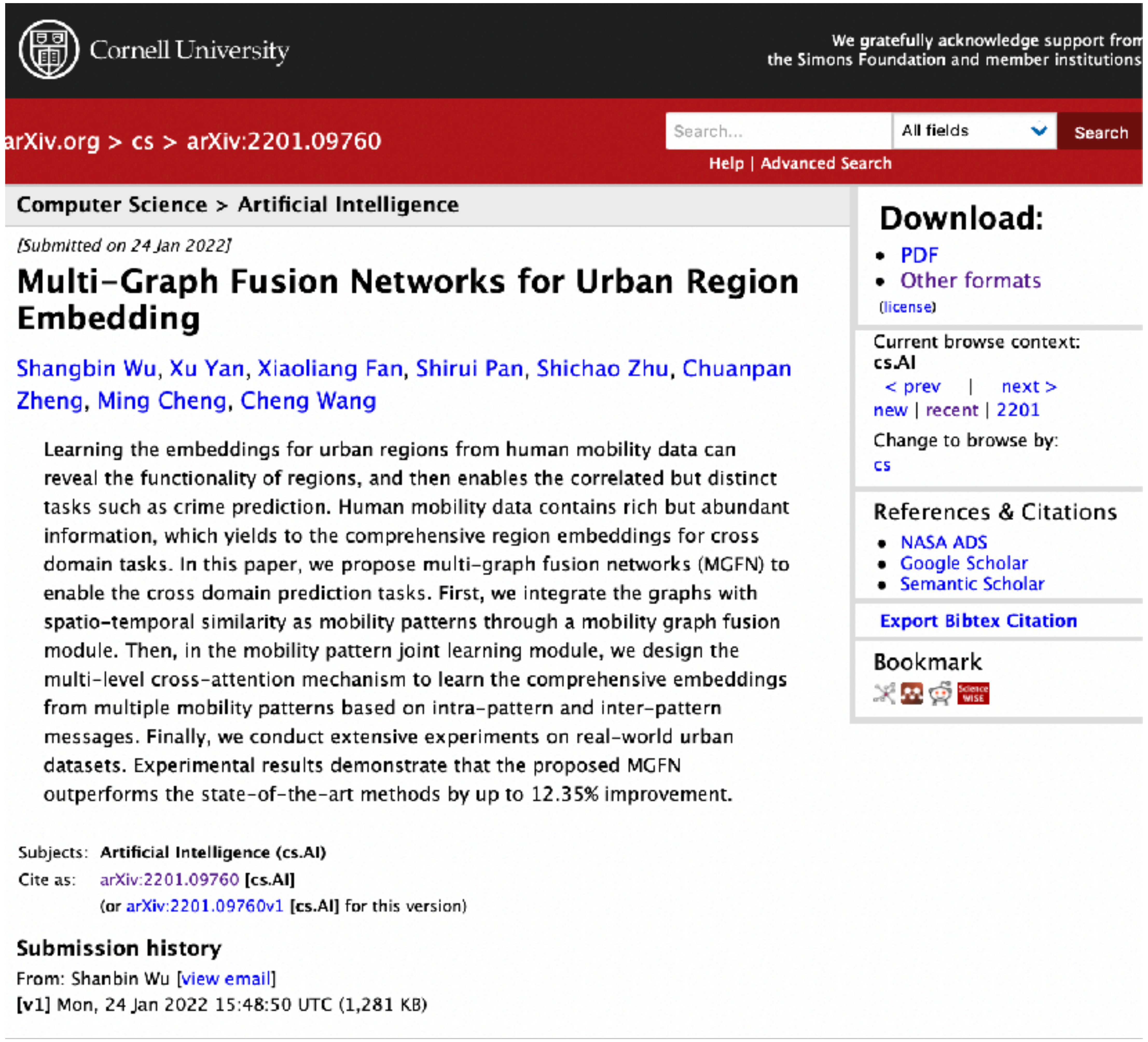
“other" format



- arXiv downloads “other" as a file without extension
- it is actually tar.gz file
- It will unpack to the current dir, so better to create a new dir
- You can find interesting points and todos there, if authors haven’t deleted them

What you can do on arXiv.org

Citations



Cornell University

We gratefully acknowledge support from the Simons Foundation and member institutions

arXiv.org > cs > arXiv:2201.09760

Search... | All fields | Search

Help | Advanced Search

Computer Science > Artificial Intelligence

[Submitted on 24 Jan 2022]

Multi-Graph Fusion Networks for Urban Region Embedding

Shangbin Wu, Xu Yan, Xiaoliang Fan, Shirui Pan, Shichao Zhu, Chuanpan Zheng, Ming Cheng, Cheng Wang

Learning the embeddings for urban regions from human mobility data can reveal the functionality of regions, and then enables the correlated but distinct tasks such as crime prediction. Human mobility data contains rich but abundant information, which yields to the comprehensive region embeddings for cross domain tasks. In this paper, we propose multi-graph fusion networks (MGFN) to enable the cross domain prediction tasks. First, we integrate the graphs with spatio-temporal similarity as mobility patterns through a mobility graph fusion module. Then, in the mobility pattern joint learning module, we design the multi-level cross-attention mechanism to learn the comprehensive embeddings from multiple mobility patterns based on intra-pattern and inter-pattern messages. Finally, we conduct extensive experiments on real-world urban datasets. Experimental results demonstrate that the proposed MGFN outperforms the state-of-the-art methods by up to 12.35% improvement.

Subjects: **Artificial Intelligence (cs.AI)**
Cite as: arXiv:2201.09760 **[cs.AI]**
(or arXiv:2201.09760v1 **[cs.AI]** for this version)

Submission history
From: Shanbin Wu [view email]
[v1] Mon, 24 Jan 2022 15:48:50 UTC (1,281 KB)

Download:
- PDF
- Other formats

(license)

Current browse context:
cs.AI
< prev | next >
new | recent | 2201
Change to browse by:
cs

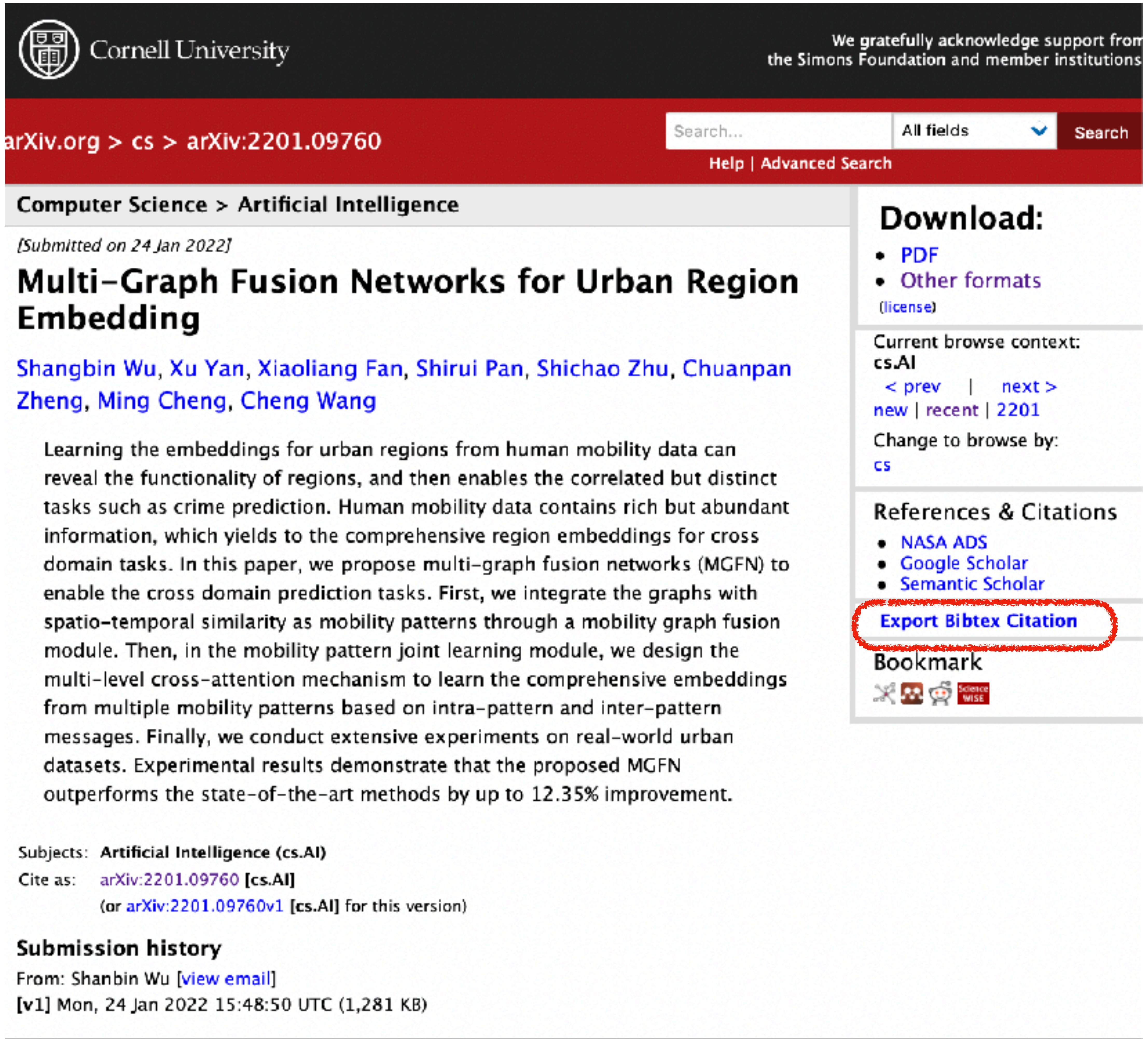
References & Citations
- NASA ADS
- Google Scholar
- Semantic Scholar

Export Bibtex Citation

Bookmark

What you can do on arXiv.org

Citations



Cornell University

We gratefully acknowledge support from the Simons Foundation and member institutions

arXiv.org > cs > arXiv:2201.09760

Search... | All fields | Search

Help | Advanced Search

Computer Science > Artificial Intelligence

[Submitted on 24 Jan 2022]

Multi-Graph Fusion Networks for Urban Region Embedding

Shangbin Wu, Xu Yan, Xiaoliang Fan, Shirui Pan, Shichao Zhu, Chuanpan Zheng, Ming Cheng, Cheng Wang

Learning the embeddings for urban regions from human mobility data can reveal the functionality of regions, and then enables the correlated but distinct tasks such as crime prediction. Human mobility data contains rich but abundant information, which yields to the comprehensive region embeddings for cross domain tasks. In this paper, we propose multi-graph fusion networks (MGFN) to enable the cross domain prediction tasks. First, we integrate the graphs with spatio-temporal similarity as mobility patterns through a mobility graph fusion module. Then, in the mobility pattern joint learning module, we design the multi-level cross-attention mechanism to learn the comprehensive embeddings from multiple mobility patterns based on intra-pattern and inter-pattern messages. Finally, we conduct extensive experiments on real-world urban datasets. Experimental results demonstrate that the proposed MGFN outperforms the state-of-the-art methods by up to 12.35% improvement.

Subjects: **Artificial Intelligence (cs.AI)**

Cite as: arXiv:2201.09760 **[cs.AI]**

(or arXiv:2201.09760v1 **[cs.AI]** for this version)

Submission history

From: Shanbin Wu [view email]

[v1] Mon, 24 Jan 2022 15:48:50 UTC (1,281 KB)

Download:

- PDF
- Other formats

(license)

Current browse context:

cs.AI

< prev | next >

new | recent | 2201

Change to browse by:

cs

References & Citations

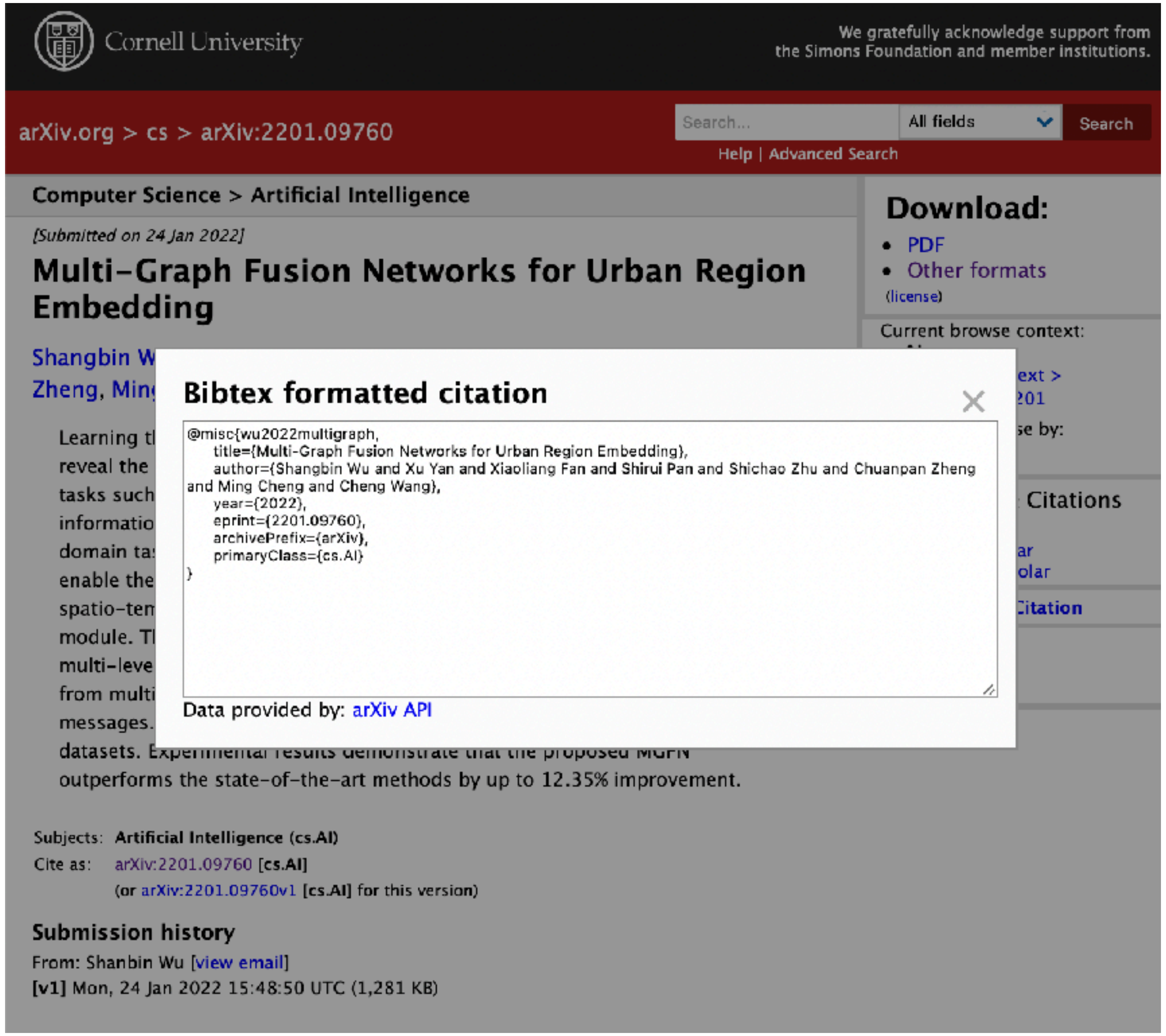
- NASA ADS
- Google Scholar
- Semantic Scholar

Export Bibtex Citation

Bookmark

What you can do on arXiv.org

Citations



Where to look for papers on arXiv

arxiv-sanity.com

Where to look for papers on arXiv

arxiv-sanity.com

- Developed by Andrew Karpathy

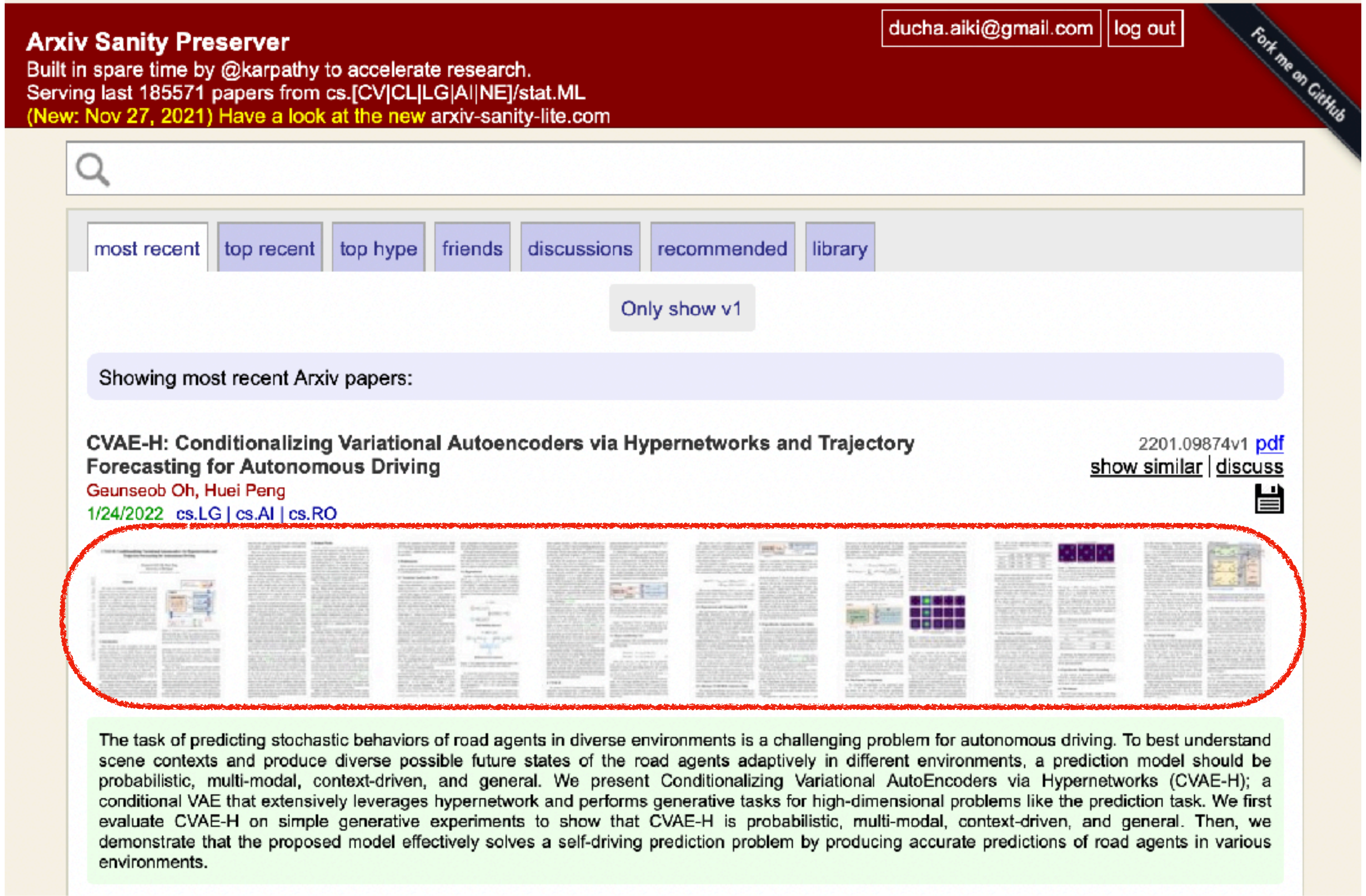
Where to look for papers on arXiv

arxiv-sanity.com

- Developed by Andrew Karpathy
- Has uber-feature: page thumbnails

Where to look for papers on arXiv

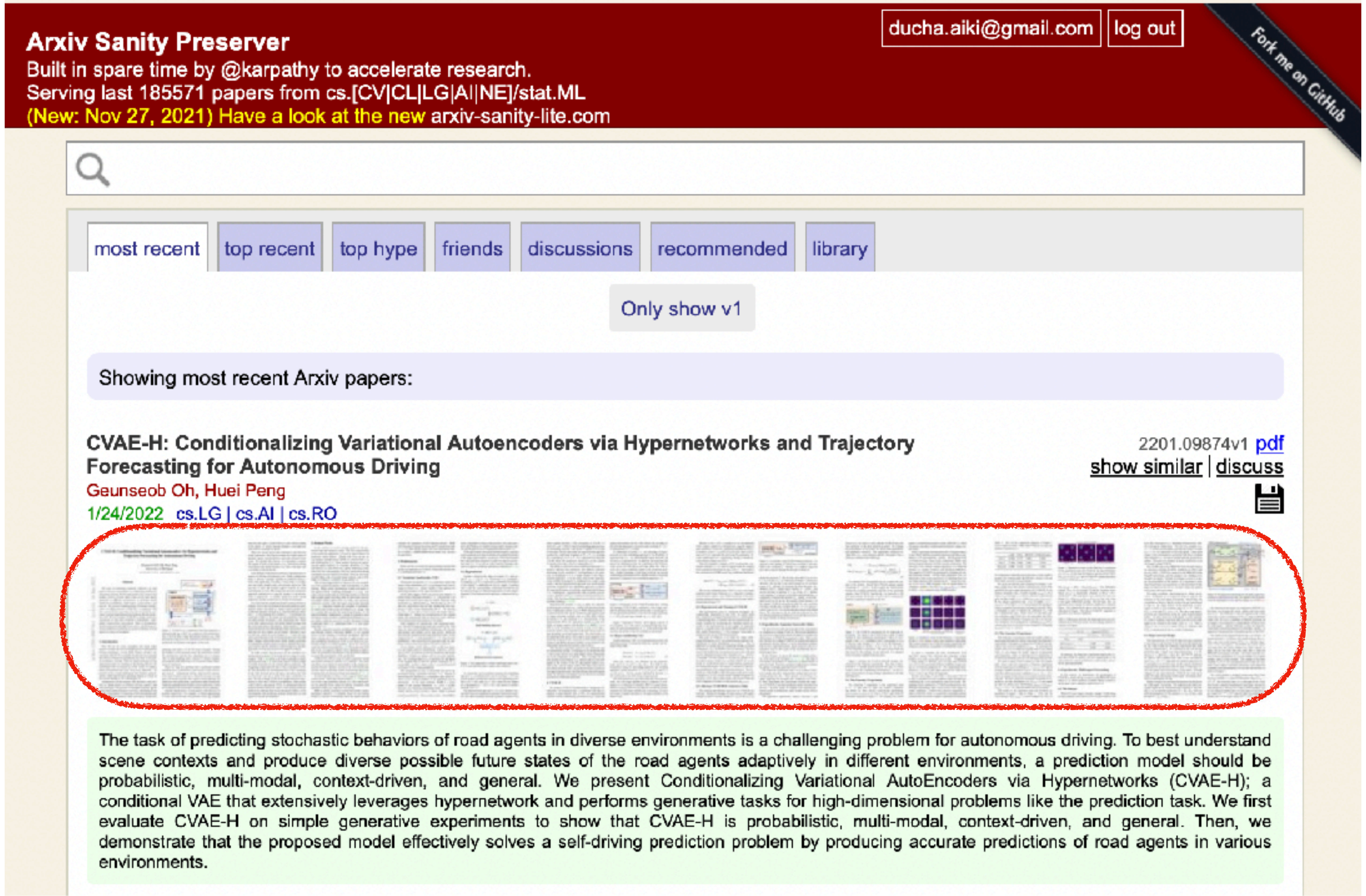
arxiv-sanity.com



- Developed by Andrew Karpathy
- Has uber-feature: page thumbnails

Where to look for papers on arXiv

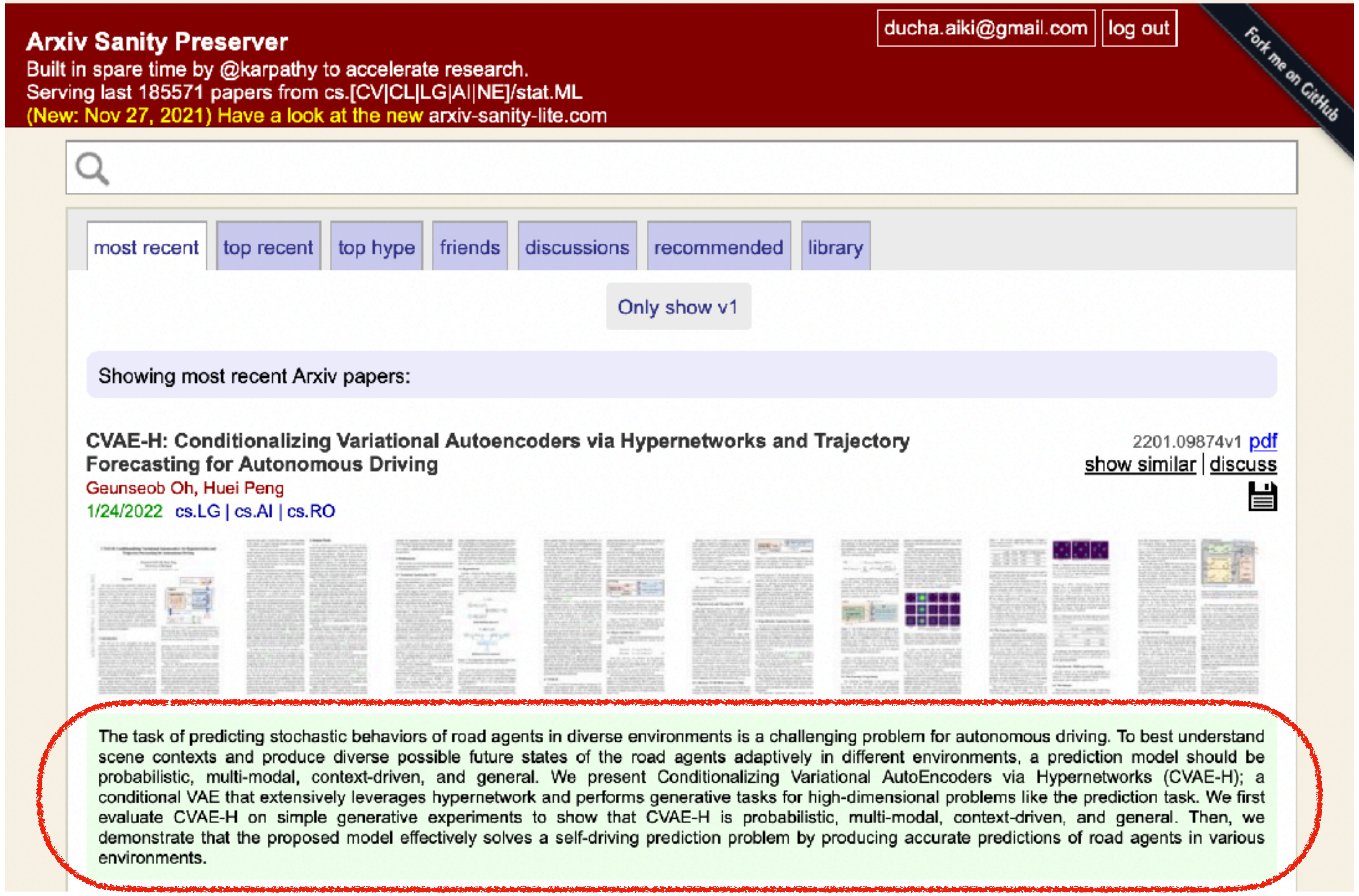
arxiv-sanity.com



- Developed by Andrew Karpathy
- Has uber-feature: page thumbnails
- Also shows paper abstract

Where to look for papers on arXiv

arxiv-sanity.com



- Developed by Andrew Karpathy
- Has uber-feature: page thumbnails
- Also shows paper abstract

Where to look for papers on arXiv

arxiv-sanity.com

Where to look for papers on arXiv

arxiv-sanity.com

Where to look for papers on arXiv

arxiv-sanity.com

- You can save paper to the library (only 500 last)

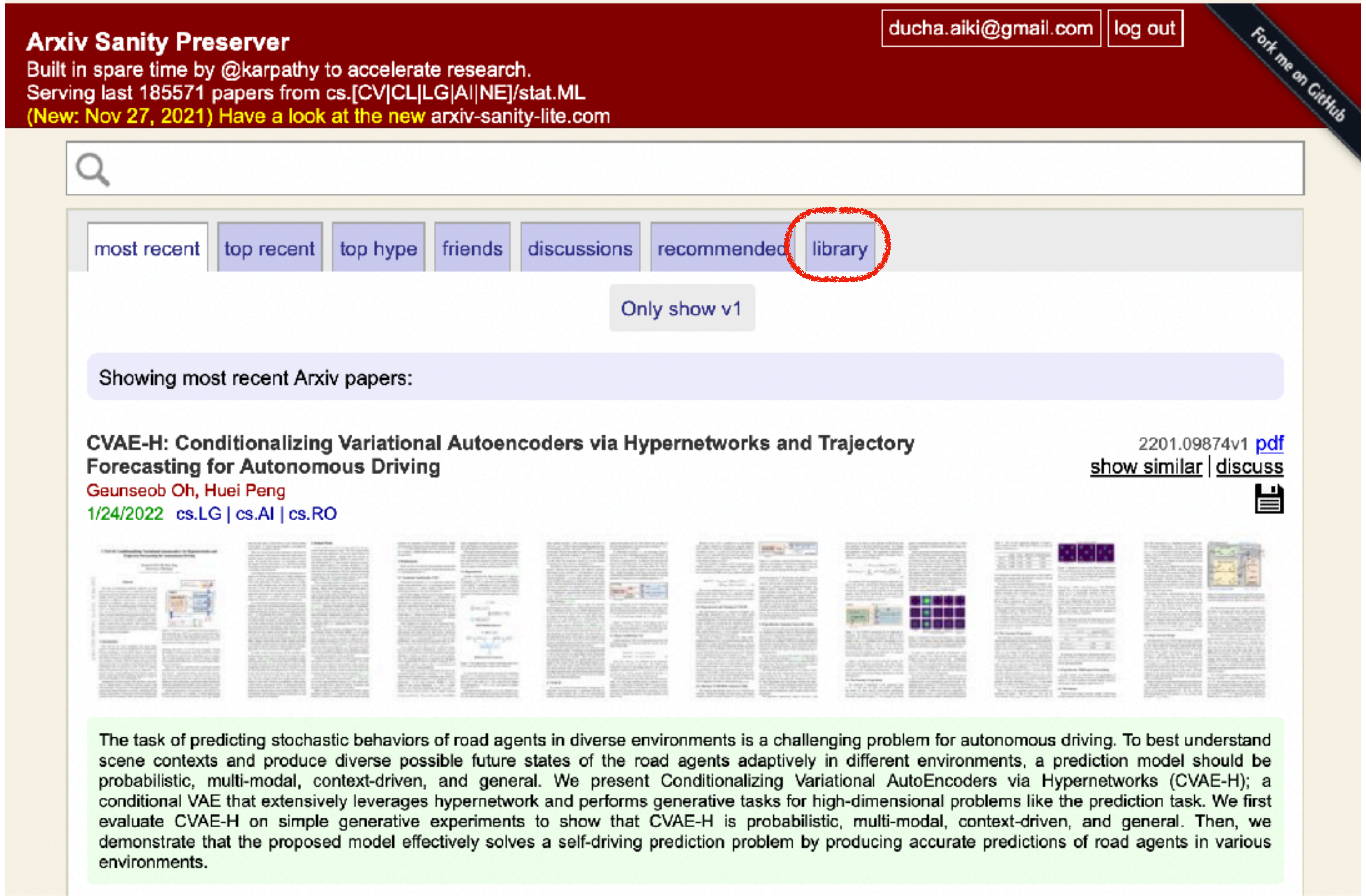
Where to look for papers on arXiv

arxiv-sanity.com

- You can save paper to the library (only 500 last)
- Based on papers you save, arxiv-sanity recommends you new papers

Where to look for papers on arXiv

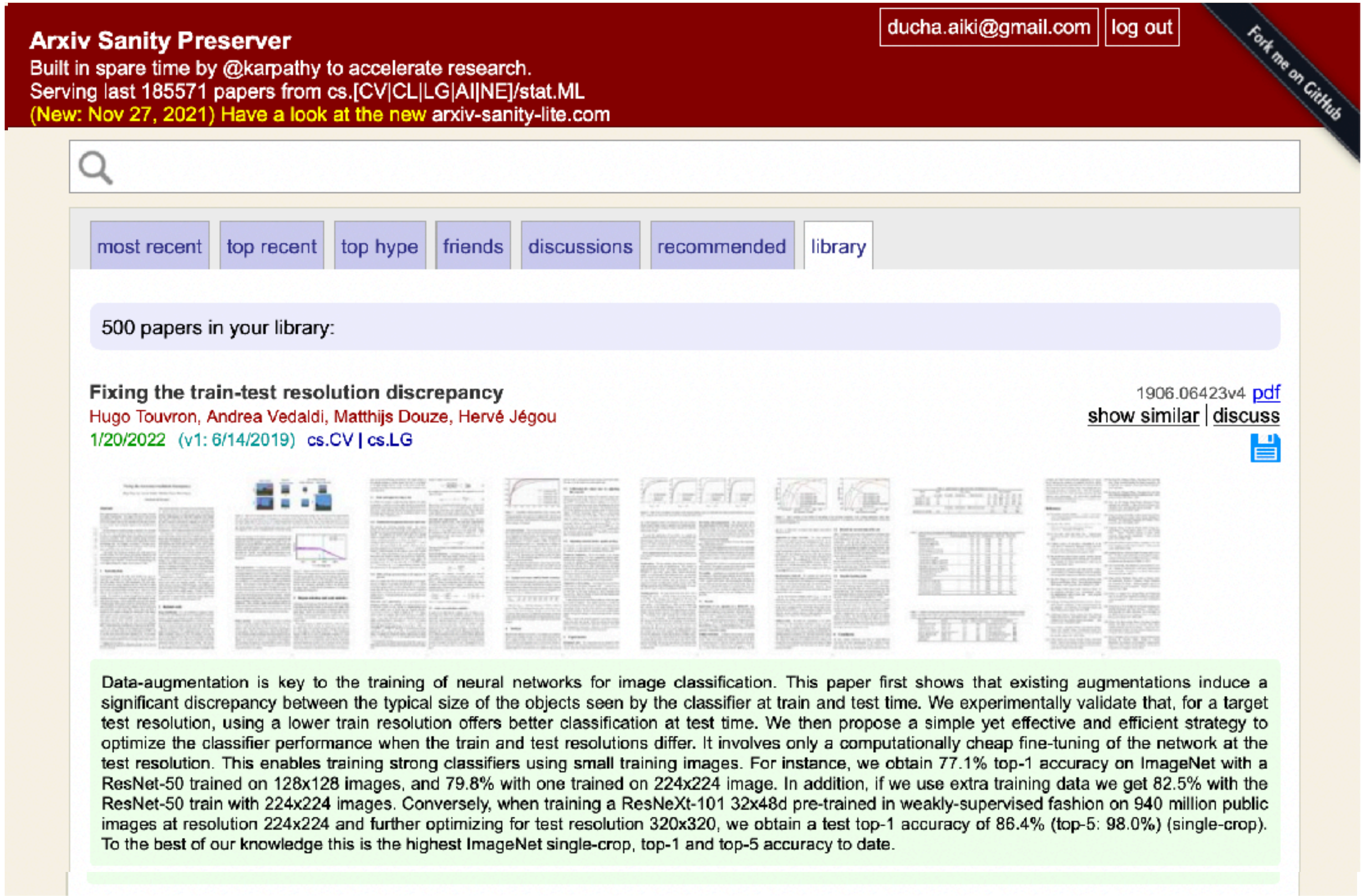
arxiv-sanity.com



- You can save paper to the library (only 500 last)
- Based on papers you save, arxiv-sanity recommends you new papers

Where to look for papers on arXiv

arxiv-sanity.com



- You can save paper to the library (only 500 last)
- Based on papers you save, arxiv-sanity recommends you new papers

Where to look for papers on arXiv

arxiv-sanity.com

- You can save paper to the library (only 500 last)
- Based on papers you save, arxiv-sanity recommends you new papers

Where to look for papers on arXiv

arxiv-sanity.com

- You can save paper to the library (only 500 last)
- Based on papers you save, arxiv-sanity recommends you new papers

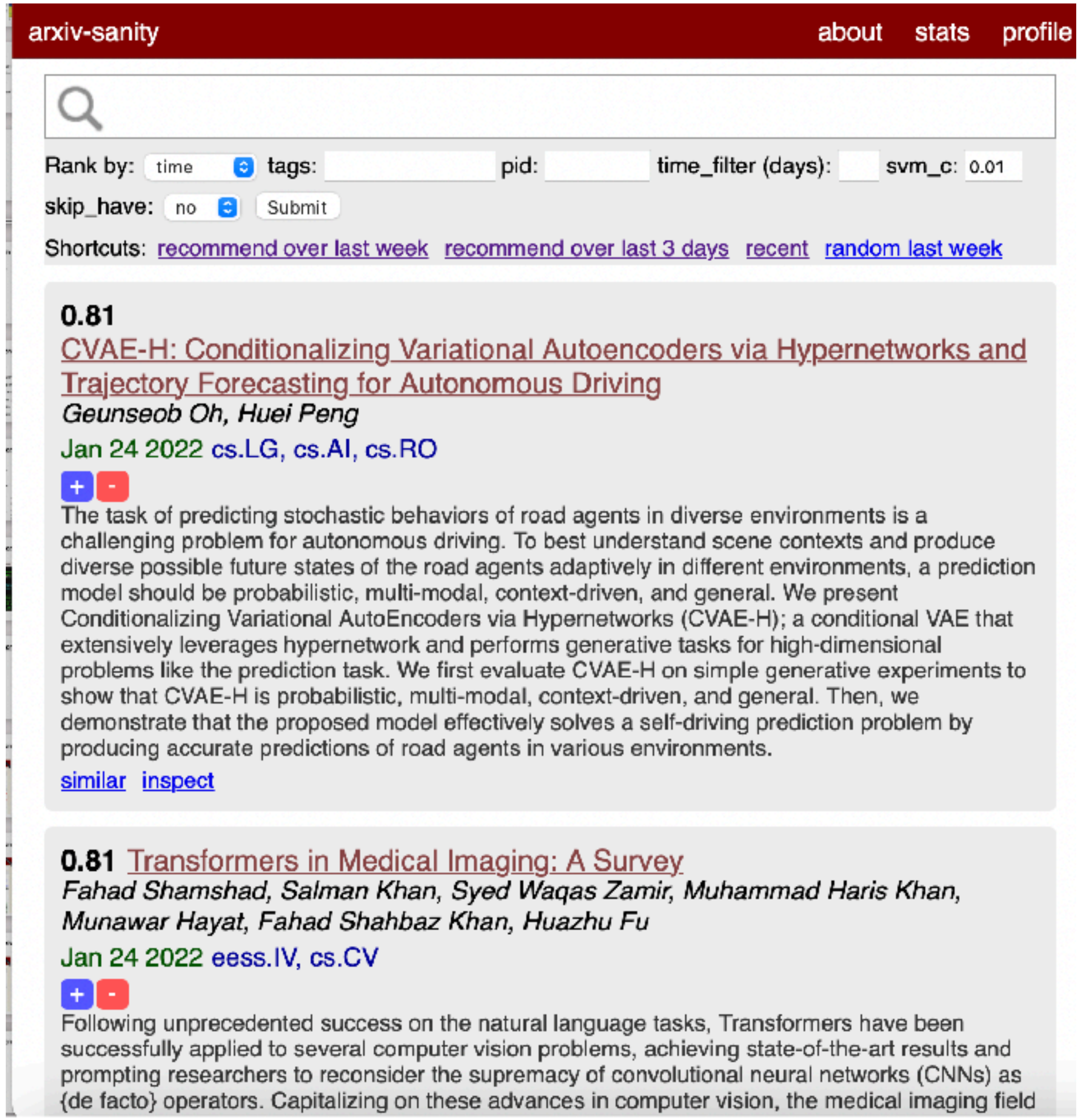
The main problem with arxiv-sanity

It is often offline

It is often offline :(

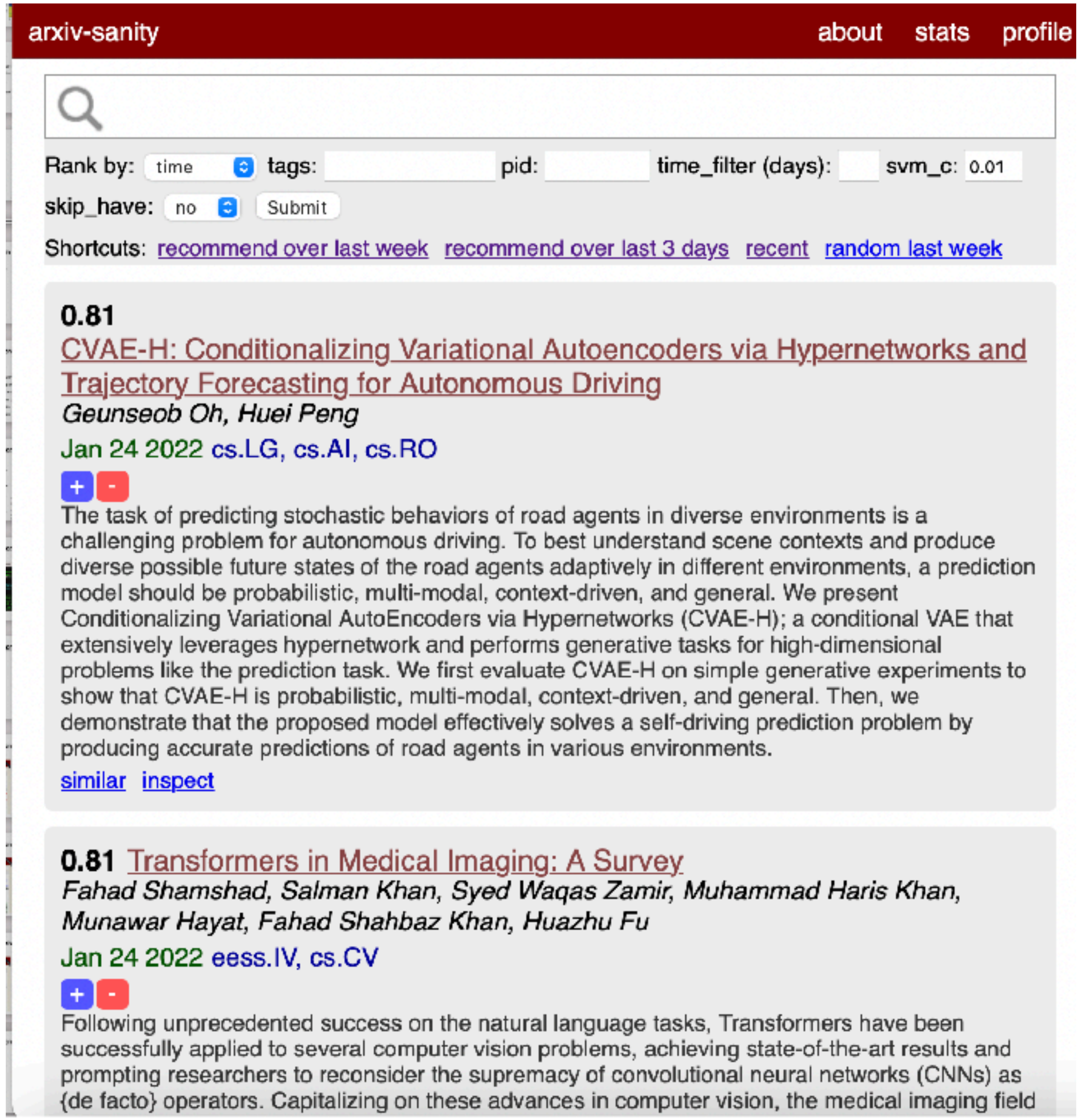
Where to look for papers on arXiv

arxiv-sanity-lite.com



Where to look for papers on arXiv

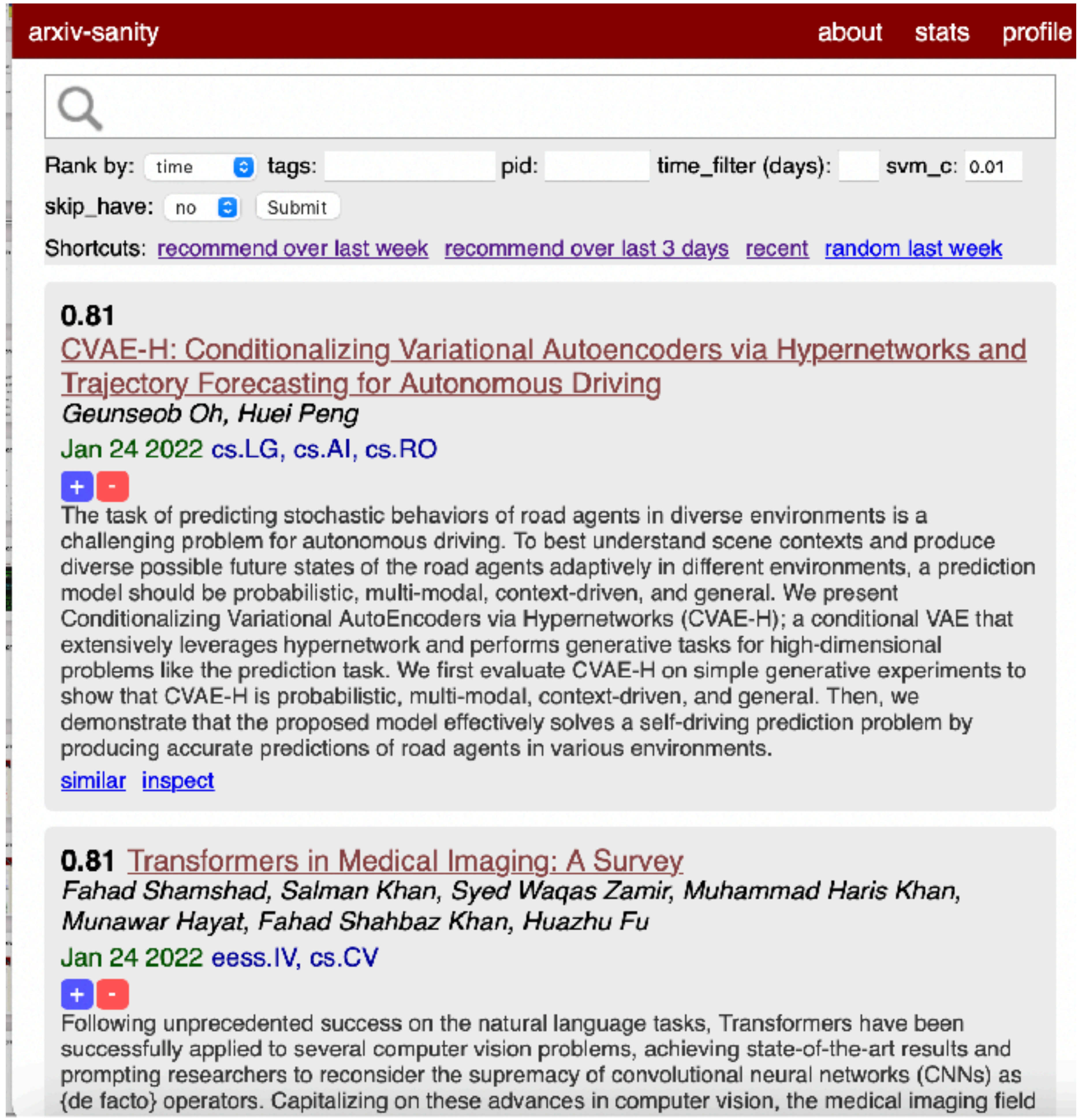
arxiv-sanity-lite.com



- Re-vamp of arxiv-sanity

Where to look for papers on arXiv

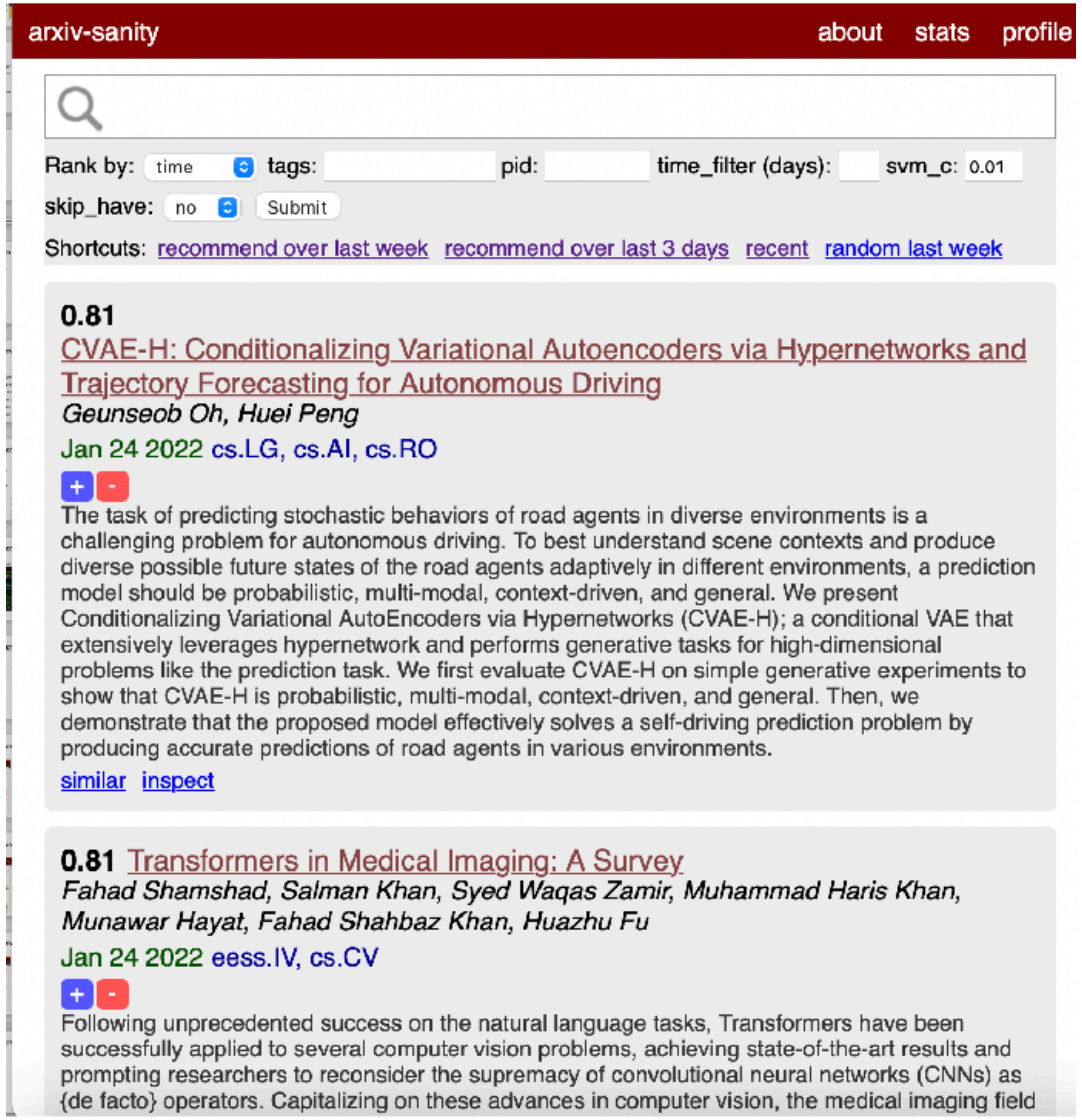
arxiv-sanity-lite.com



- Re-vamp of arxiv-sanity
- No page thumbnails

Where to look for papers on arXiv

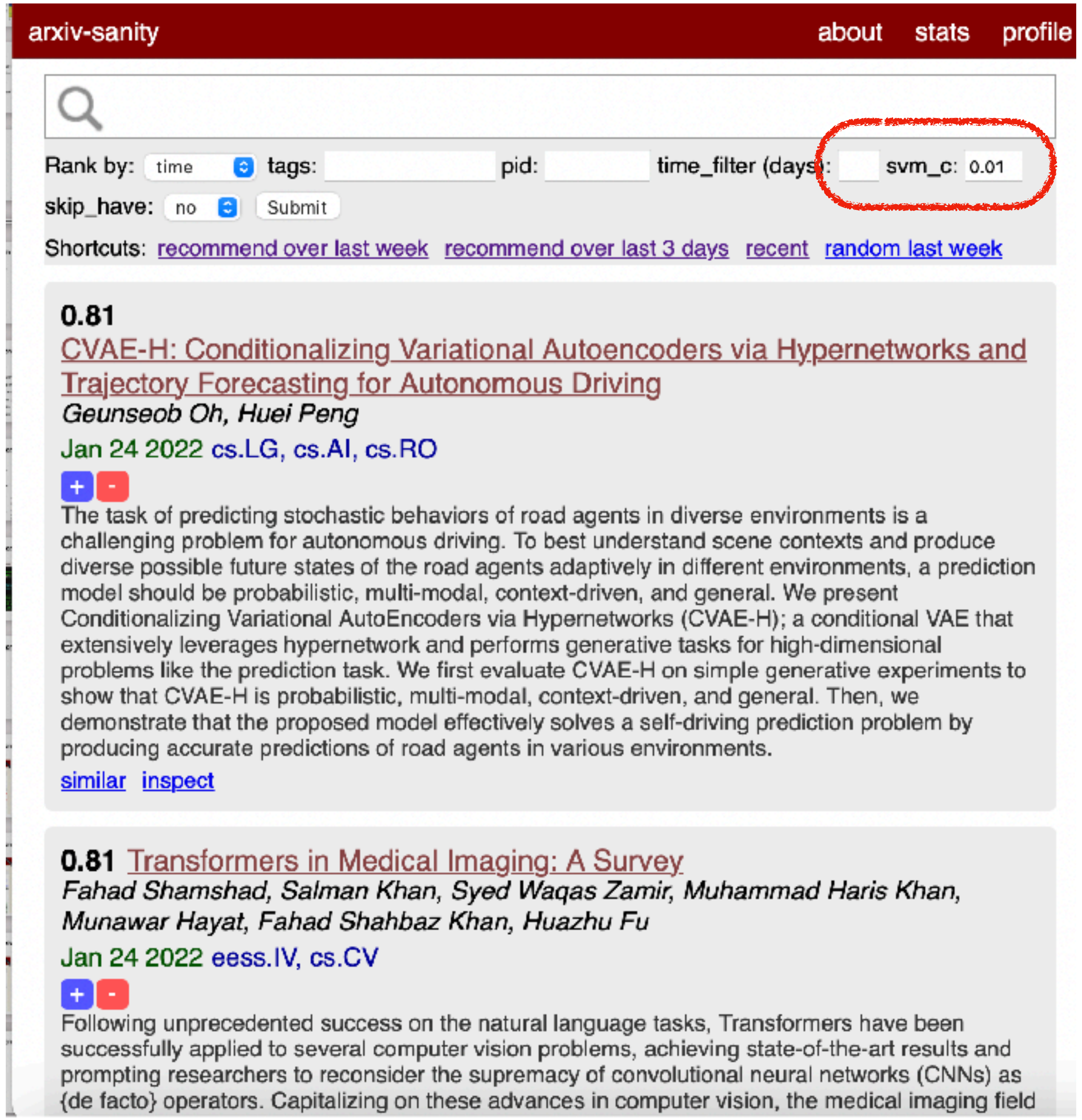
arxiv-sanity-lite.com



- Re-vamp of arxiv-sanity
- No page thumbnails
- You can control recommendations from tf-idf SVM

Where to look for papers on arXiv

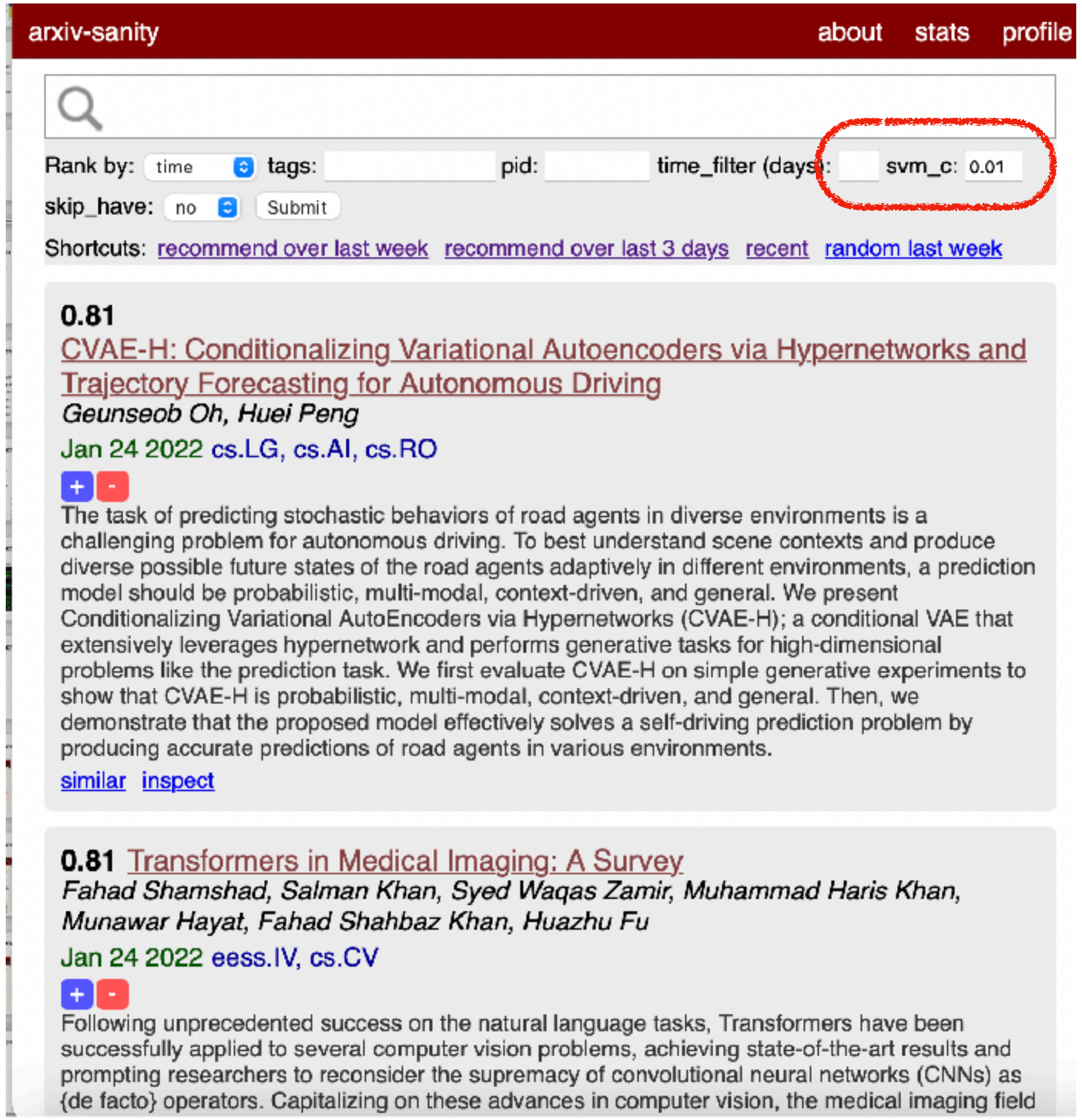
arxiv-sanity-lite.com



- Re-vamp of arxiv-sanity
- No page thumbnails
- You can control recommendations from tf-idf SVM

Where to look for papers on arXiv

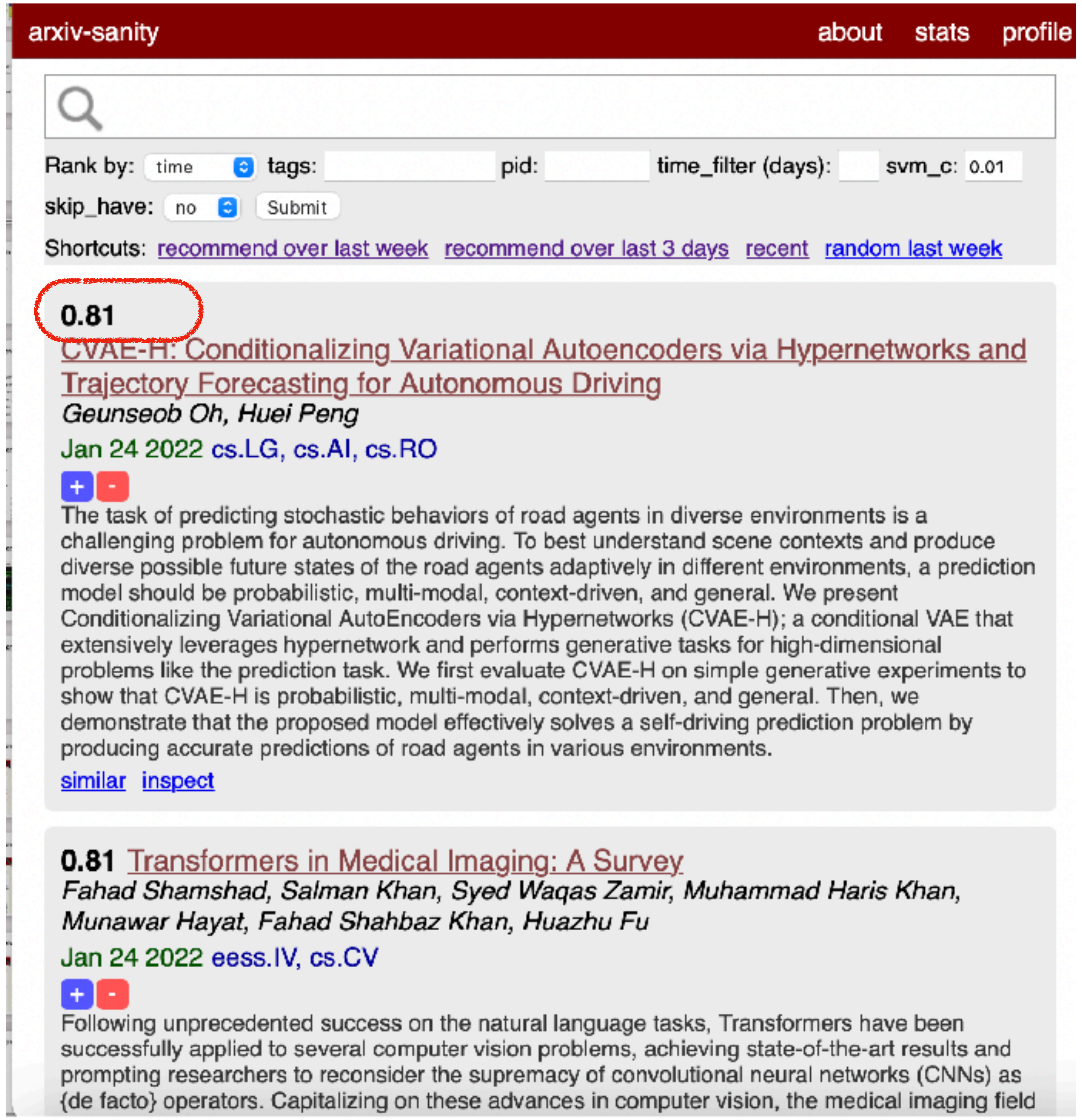
arxiv-sanity-lite.com



- Re-vamp of arxiv-sanity
- No page thumbnails
- You can control recommendations from tf-idf SVM
- “Relevancy score”

Where to look for papers on arXiv

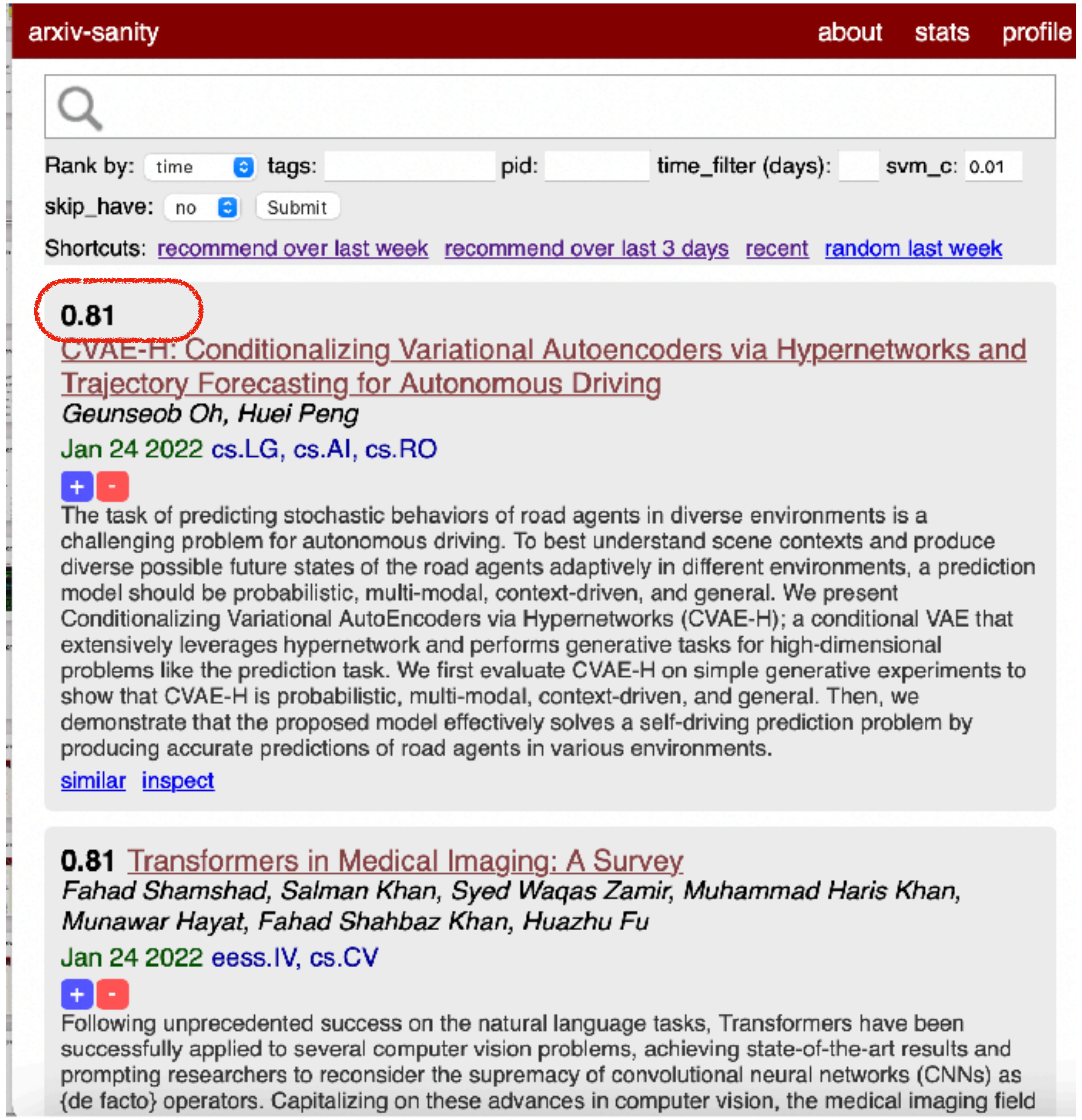
arxiv-sanity-lite.com



- Re-vamp of arxiv-sanity
- No page thumbnails
- You can control recommendations from tf-idf SVM
- “Relevancy score”

Where to look for papers on arXiv

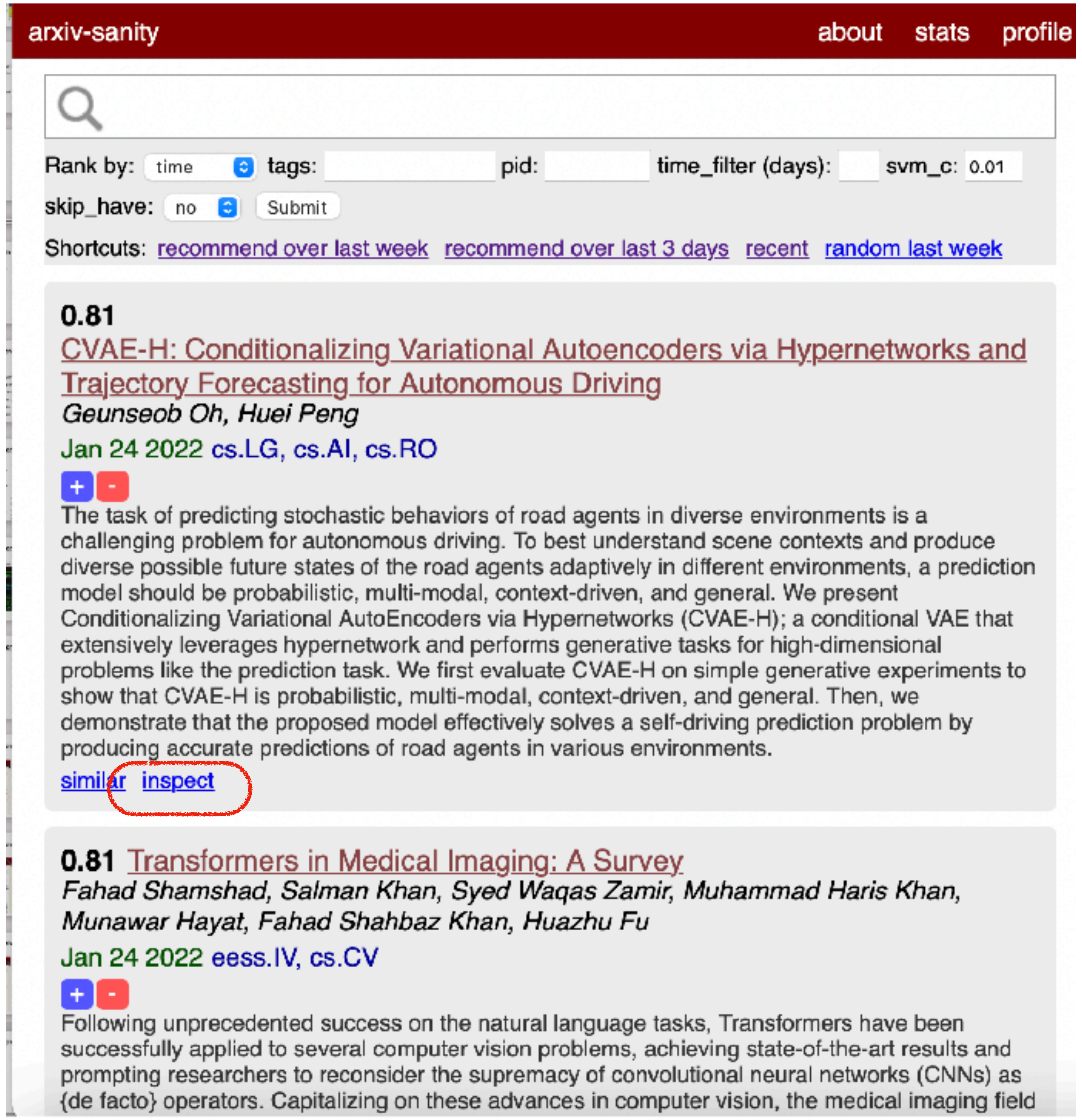
arxiv-sanity-lite.com



- Re-vamp of arxiv-sanity
- No page thumbnails
- You can control recommendations from tf-idf SVM
- “Relevancy score”
- “Inspect" the terms used for relevancy score calculation

Where to look for papers on arXiv

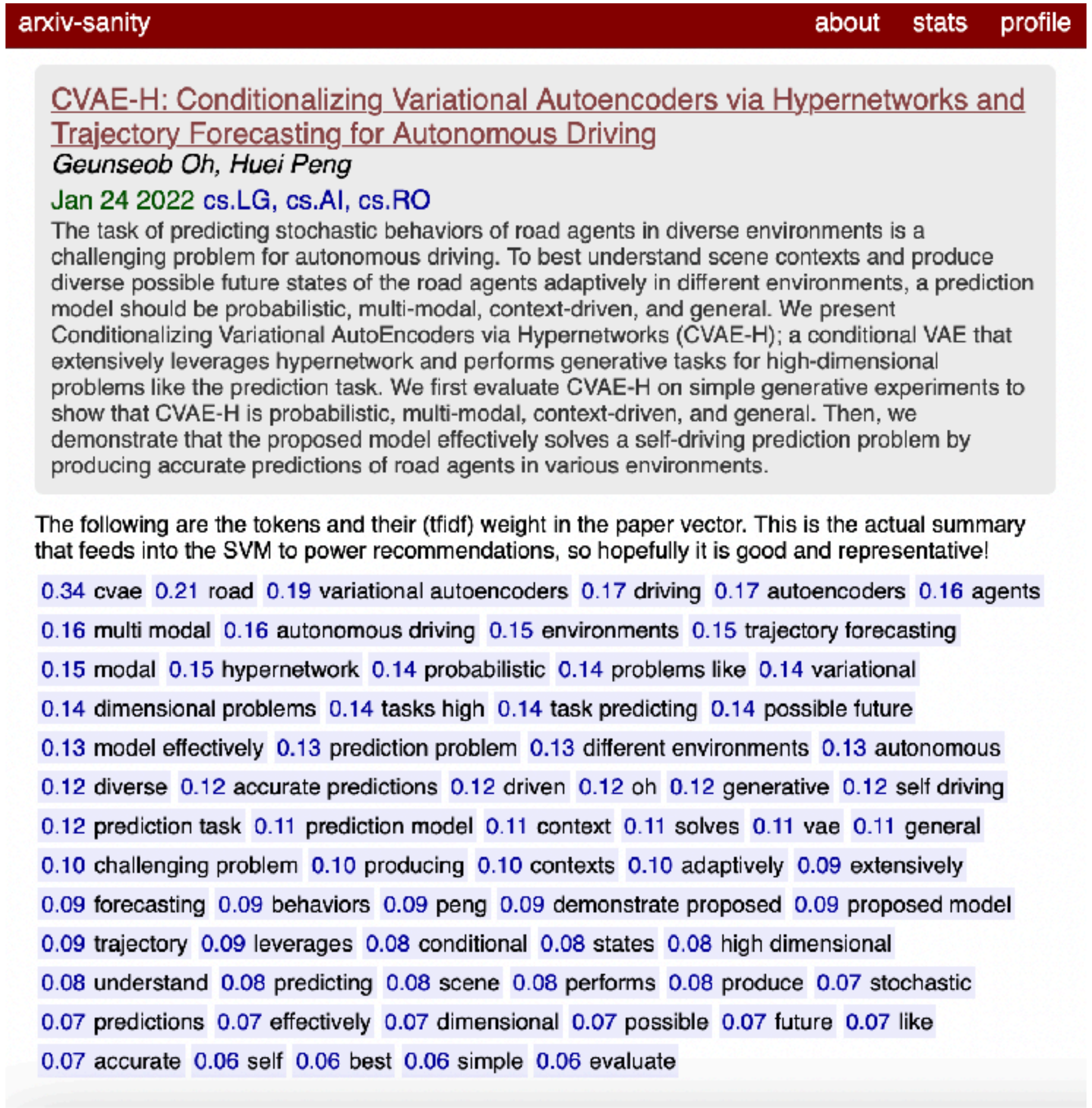
arxiv-sanity-lite.com



- Re-vamp of arxiv-sanity
- No page thumbnails
- You can control recommendations from tf-idf SVM
- “Relevancy score”
- “Inspect" the terms used for relevancy score calculation

Where to look for papers on arXiv

arxiv-sanity-lite.com


arxiv-sanity about stats profile

CVAE-H: Conditionalizing Variational Autoencoders via Hypernetworks and Trajectory Forecasting for Autonomous Driving

Geunseob Oh, Huei Peng

Jan 24 2022 cs.LG, cs.AI, cs.RO

The task of predicting stochastic behaviors of road agents in diverse environments is a challenging problem for autonomous driving. To best understand scene contexts and produce diverse possible future states of the road agents adaptively in different environments, a prediction model should be probabilistic, multi-modal, context-driven, and general. We present Conditionalizing Variational AutoEncoders via Hypernetworks (CVAE-H); a conditional VAE that extensively leverages hypernetwork and performs generative tasks for high-dimensional problems like the prediction task. We first evaluate CVAE-H on simple generative experiments to show that CVAE-H is probabilistic, multi-modal, context-driven, and general. Then, we demonstrate that the proposed model effectively solves a self-driving prediction problem by producing accurate predictions of road agents in various environments.

The following are the tokens and their (tfidf) weight in the paper vector. This is the actual summary that feeds into the SVM to power recommendations, so hopefully it is good and representative!

0.34 cvae 0.21 road 0.19 variational autoencoders 0.17 driving 0.17 autoencoders 0.16 agents
0.16 multi modal 0.16 autonomous driving 0.15 environments 0.15 trajectory forecasting
0.15 modal 0.15 hypernetwork 0.14 probabilistic 0.14 problems like 0.14 variational
0.14 dimensional problems 0.14 tasks high 0.14 task predicting 0.14 possible future
0.13 model effectively 0.13 prediction problem 0.13 different environments 0.13 autonomous
0.12 diverse 0.12 accurate predictions 0.12 driven 0.12 oh 0.12 generative 0.12 self driving
0.12 prediction task 0.11 prediction model 0.11 context 0.11 solves 0.11 vae 0.11 general
0.10 challenging problem 0.10 producing 0.10 contexts 0.10 adaptively 0.09 extensively
0.09 forecasting 0.09 behaviors 0.09 peng 0.09 demonstrate proposed 0.09 proposed model
0.09 trajectory 0.09 leverages 0.08 conditional 0.08 states 0.08 high dimensional
0.08 understand 0.08 predicting 0.08 scene 0.08 performs 0.08 produce 0.07 stochastic
0.07 predictions 0.07 effectively 0.07 dimensional 0.07 possible 0.07 future 0.07 like
0.07 accurate 0.06 self 0.06 best 0.06 simple 0.06 evaluate

Where to look for papers on arXiv

arxiv-sanity-lite.com

2.00 Learning and Crafting for the Wide Multiple Baseline Stereo
Dmytro Mishkin
Dec 22 2021 cs.CV
+ -
This thesis introduces the wide multiple baseline stereo (WxBS) problem. WxBS, a generalization of the standard wide baseline stereo problem, considers the matching of images that simultaneously differ in more than one image acquisition factor such as viewpoint, illumination, sensor type, or where object appearance changes significantly, e.g., over time. A new dataset with the ground truth, evaluation metric and baselines has been introduced. The thesis presents the following improvements of the WxBS pipeline. (i) A loss function, called HardNeg, for learning a local image descriptor that relies on hard negative mining within a mini-batch and on the maximization of the distance between the closest positive and the closest negative patches. (ii) The descriptor trained with the HardNeg loss, called HardNet, is compact and shows state-of-the-art performance in standard matching, patch verification and retrieval benchmarks. (iii) A method for learning the affine shape, orientation, and potentially other parameters related to geometric and appearance properties of local features. (iv) A tentative correspondences generation strategy which generalizes the standard first to second closest distance ratio is presented. The selection strategy, which shows performance superior to the standard method, is applicable to either hard-engineered descriptors like SIFT, LIOP, and MROGH or deeply learned like HardNet. (v) A feedback loop is introduced for the two-view matching problem, resulting in MODS -- matching with on-demand view synthesis -- algorithm. MODS is an algorithm that handles a viewing angle difference even larger than the previous state-of-the-art ASIFT algorithm, without a significant increase of computational cost over "standard" wide and narrow baseline approaches. Last, but not least, a comprehensive benchmark for local features and robust estimation algorithms is introduced.
similar inspect

Where to look for papers on arXiv

arxiv-sanity-lite.com

2.00 Learning and Crafting for the Wide Multiple Baseline Stereo
Dmytro Mishkin
Dec 22 2021 cs.CV
+ -
This thesis introduces the wide multiple baseline stereo (WxBS) problem. WxBS, a generalization of the standard wide baseline stereo problem, considers the matching of images that simultaneously differ in more than one image acquisition factor such as viewpoint, illumination, sensor type, or where object appearance changes significantly, e.g., over time. A new dataset with the ground truth, evaluation metric and baselines has been introduced. The thesis presents the following improvements of the WxBS pipeline. (i) A loss function, called HardNeg, for learning a local image descriptor that relies on hard negative mining within a mini-batch and on the maximization of the distance between the closest positive and the closest negative patches. (ii) The descriptor trained with the HardNeg loss, called HardNet, is compact and shows state-of-the-art performance in standard matching, patch verification and retrieval benchmarks. (iii) A method for learning the affine shape, orientation, and potentially other parameters related to geometric and appearance properties of local features. (iv) A tentative correspondences generation strategy which generalizes the standard first to second closest distance ratio is presented. The selection strategy, which shows performance superior to the standard method, is applicable to either hard-engineered descriptors like SIFT, LIOP, and MROGH or deeply learned like HardNet. (v) A feedback loop is introduced for the two-view matching problem, resulting in MODS -- matching with on-demand view synthesis -- algorithm. MODS is an algorithm that handles a viewing angle difference even larger than the previous state-of-the-art ASIFT algorithm, without a significant increase of computational cost over "standard" wide and narrow baseline approaches. Last, but not least, a comprehensive benchmark for local features and robust estimation algorithms is introduced.
similar inspect

- Similarity search by tf-idf

Where to look for papers on arXiv

arxiv-sanity-lite.com

2.00 Learning and Crafting for the Wide Multiple Baseline Stereo
Dmytro Mishkin
Dec 22 2021 cs.CV
+ -
This thesis introduces the wide multiple baseline stereo (WxBS) problem. WxBS, a generalization of the standard wide baseline stereo problem, considers the matching of images that simultaneously differ in more than one image acquisition factor such as viewpoint, illumination, sensor type, or where object appearance changes significantly, e.g., over time. A new dataset with the ground truth, evaluation metric and baselines has been introduced. The thesis presents the following improvements of the WxBS pipeline. (i) A loss function, called HardNeg, for learning a local image descriptor that relies on hard negative mining within a mini-batch and on the maximization of the distance between the closest positive and the closest negative patches. (ii) The descriptor trained with the HardNeg loss, called HardNet, is compact and shows state-of-the-art performance in standard matching, patch verification and retrieval benchmarks. (iii) A method for learning the affine shape, orientation, and potentially other parameters related to geometric and appearance properties of local features. (iv) A tentative correspondences generation strategy which generalizes the standard first to second closest distance ratio is presented. The selection strategy, which shows performance superior to the standard method, is applicable to either hard-engineered descriptors like SIFT, LIOP, and MROGH or deeply learned like HardNet. (v) A feedback loop is introduced for the two-view matching problem, resulting in MODS -- matching with on-demand view synthesis -- algorithm. MODS is an algorithm that handles a viewing angle difference even larger than the previous state-of-the-art ASIFT algorithm, without a significant increase of computational cost over "standard" wide and narrow baseline approaches. Last, but not least, a comprehensive benchmark for local features and robust estimation algorithms is introduced.
similar inspect

- Similarity search by tf-idf

Where to look for papers on arXiv

arxiv-sanity-lite.com

2.00 Learning and Crafting for the Wide Multiple Baseline Stereo
Dmytro Mishkin
Dec 22 2021 cs.CV
+ -
This thesis introduces the wide multiple baseline stereo (WxBS) problem. WxBS, a generalization of the standard wide baseline stereo problem, considers the matching of images that simultaneously differ in more than one image acquisition factor such as viewpoint, illumination, sensor type, or where object appearance changes significantly, e.g., over time. A new dataset with the ground truth, evaluation metric and baselines has been introduced. The thesis presents the following improvements of the WxBS pipeline. (i) A loss function, called HardNeg, for learning a local image descriptor that relies on hard negative mining within a mini-batch and on the maximization of the distance between the closest positive and the closest negative patches. (ii) The descriptor trained with the HardNeg loss, called HardNet, is compact and shows state-of-the-art performance in standard matching, patch verification and retrieval benchmarks. (iii) A method for learning the affine shape, orientation, and potentially other parameters related to geometric and appearance properties of local features. (iv) A tentative correspondences generation strategy which generalizes the standard first to second closest distance ratio is presented. The selection strategy, which shows performance superior to the standard method, is applicable to either hard-engineered descriptors like SIFT, LIOP, and MROGH or deeply learned like HardNet. (v) A feedback loop is introduced for the two-view matching problem, resulting in MODS -- matching with on-demand view synthesis -- algorithm. MODS is an algorithm that handles a viewing angle difference even larger than the previous state-of-the-art ASIFT algorithm, without a significant increase of computational cost over "standard" wide and narrow baseline approaches. Last, but not least, a comprehensive benchmark for local features and robust estimation algorithms is introduced.
similar inspect

- Similarity search by tf-idf

-69.66 Digging Into Self-Supervised Learning of Feature Descriptors
Iaroslav Melekhov, Zakaria Laskar, Xiaotian Li, Shuzhe Wang, Juho Kannala
Oct 10 2021 cs.CV
+ -
Fully-supervised CNN-based approaches for learning local image descriptors have shown remarkable results in a wide range of geometric tasks. However, most of them require per-pixel ground-truth keypoint correspondence data which is difficult to acquire at scale. To address this challenge, recent weakly- and self-supervised methods can learn feature descriptors from relative camera poses or using only synthetic rigid transformations such as homographies. In this work, we focus on understanding the limitations of existing self-supervised approaches and propose a set of improvements that combined lead to powerful feature descriptors. We show that increasing the search space from in-pair to in-batch for hard negative mining brings consistent improvement. To enhance the discriminativeness of feature descriptors, we propose a coarse-to-fine method for mining local hard negatives from a wider search space by using global visual image descriptors. We demonstrate that a combination of synthetic homography transformation, color augmentation, and photorealistic image stylization produces useful representations that are viewpoint and illumination invariant. The feature descriptors learned by the proposed approach perform competitively and surpass their fully- and weakly-supervised counterparts on various geometric benchmarks such as image-based localization, sparse feature matching, and image retrieval.
similar inspect

-72.01 SIFT Matching by Context Exposed
Fabio Bellavia
Aug 27 2021 cs.CV
+ -
This paper investigates how to step up local image descriptor matching by exploiting matching context information. Two main contexts are identified, originated respectively from the descriptor space and from the keypoint space. The former is generally used to design the actual matching strategy while the latter to filter matches according to the local spatial consistency. On this basis, a new matching strategy and a novel local spatial filter, named respectively blob matching and Delaunay Triangulation Matching (DTM) are devised. Blob matching provides a general matching framework by merging together several strategies, including rank-based pre-filtering as well as many-to-many and symmetric matching, enabling to achieve a global improvement upon each individual strategy. DTM alternates between Delaunay triangulation contractions and expansions to figure out and adjust keypoint neighborhood consistency. Experimental evaluation shows that DTM is comparable or better than the state-of-the-art in terms of matching accuracy and robustness. Evaluation is carried out according to a new benchmark devised for analyzing the matching pipeline in terms of correct correspondences on both planar and non-planar scenes, including several state-of-the-art methods as well as the common SIFT matching approach for reference. This evaluation can be of assistance for future research in this field.
similar inspect

How to read arXiv papers

Coarse-to-fine

How to read arXiv papers

Coarse-to-fine

- I skim through arxiv-sanity list (or arXiv.org CS.CV) almost daily

How to read arXiv papers

Coarse-to-fine

- I skim through arxiv-sanity list (or arXiv.org CS.CV) almost daily
- Check all titles

How to read arXiv papers

Coarse-to-fine

- I skim through arxiv-sanity list (or arXiv.org CS.CV) almost daily
- Check all titles
- Open in the new tab to check the abstract

How to read arXiv papers

Coarse-to-fine

- I skim through arxiv-sanity list (or arXiv.org CS.CV) almost daily
- Check all titles
- Open in the new tab to check the abstract
- Based on abstract open from 1 to 10 pdfs

Should read this paper?

Questions I ask myself (1)

Should read this paper?

Questions I ask myself (1)

- My areas? Yes → read.

Should read this paper?

Questions I ask myself (1)

- My areas? Yes → read.
- Opens a new (sub-)area of research? Yes → read. E.g. GAN in 2014, NERF

Should read this paper?

Questions I ask myself (1)

- My areas? Yes → read.
- Opens a new (sub-)area of research? Yes → read. E.g. GAN in 2014, NERF
- From the crowded area? Yes → skip. E.g GANs now

Should read this paper?

Questions I ask myself (1)

- My areas? Yes → read.
- Opens a new (sub-)area of research? Yes → read. E.g. GAN in 2014, NERF
- From the crowded area? Yes → skip. E.g GANs now
- Dataset or large-scale benchmark paper? Yes → read

Should read this paper?

Questions I ask myself (1)

- My areas? Yes → read.
- Opens a new (sub-)area of research? Yes → read. E.g. GAN in 2014, NERF
- From the crowded area? Yes → skip. E.g GANs now
- Dataset or large-scale benchmark paper? Yes → read
- Simple baseline? Yes → read, regadless of the area.

Should read this paper?

Questions I ask myself (1)

- My areas? Yes → read.
- Opens a new (sub-)area of research? Yes → read. E.g. GAN in 2014, NERF
- From the crowded area? Yes → skip. E.g GANs now
- Dataset or large-scale benchmark paper? Yes → read
- Simple baseline? Yes → read, regadless of the area.
- About understanding some aspect of machine learning? Yes → read if have time. E.g. padding, double descent, over-parametrization

Should read this paper?

Questions I ask myself (2)

Should read this paper?

Questions I ask myself (2)

- Theory paper? Yes → skip

Should read this paper?

Questions I ask myself (2)

- Theory paper? Yes → skip
- Relevant for me as a user? Yes → read if have time. E.g. new non-linearity, optimizer, etc.

Should read this paper?

Questions I ask myself (2)

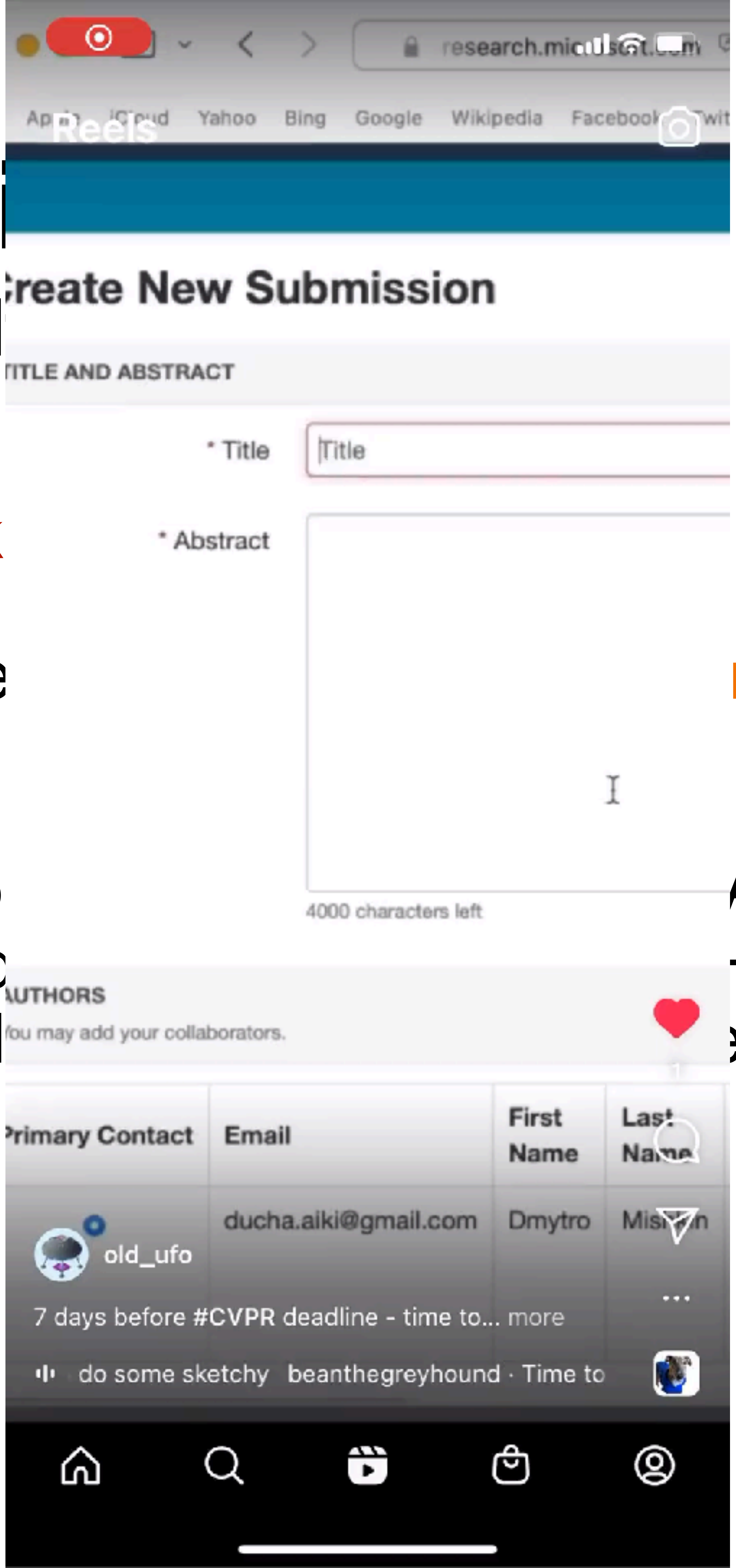
- Theory paper? Yes → skip

- Relevant for me as a user? Yes → read if have time. E.g. new non-linearity, optimizer, etc.

- Isn't the title over-keyworded and so on? E.g. "Adaptive DropBlock Enhanced Generative Adversarial Networks for Hyperspectral Image Classification”. This kind of things are hard to verbalize.

Should read thi

Questions I ask mysel

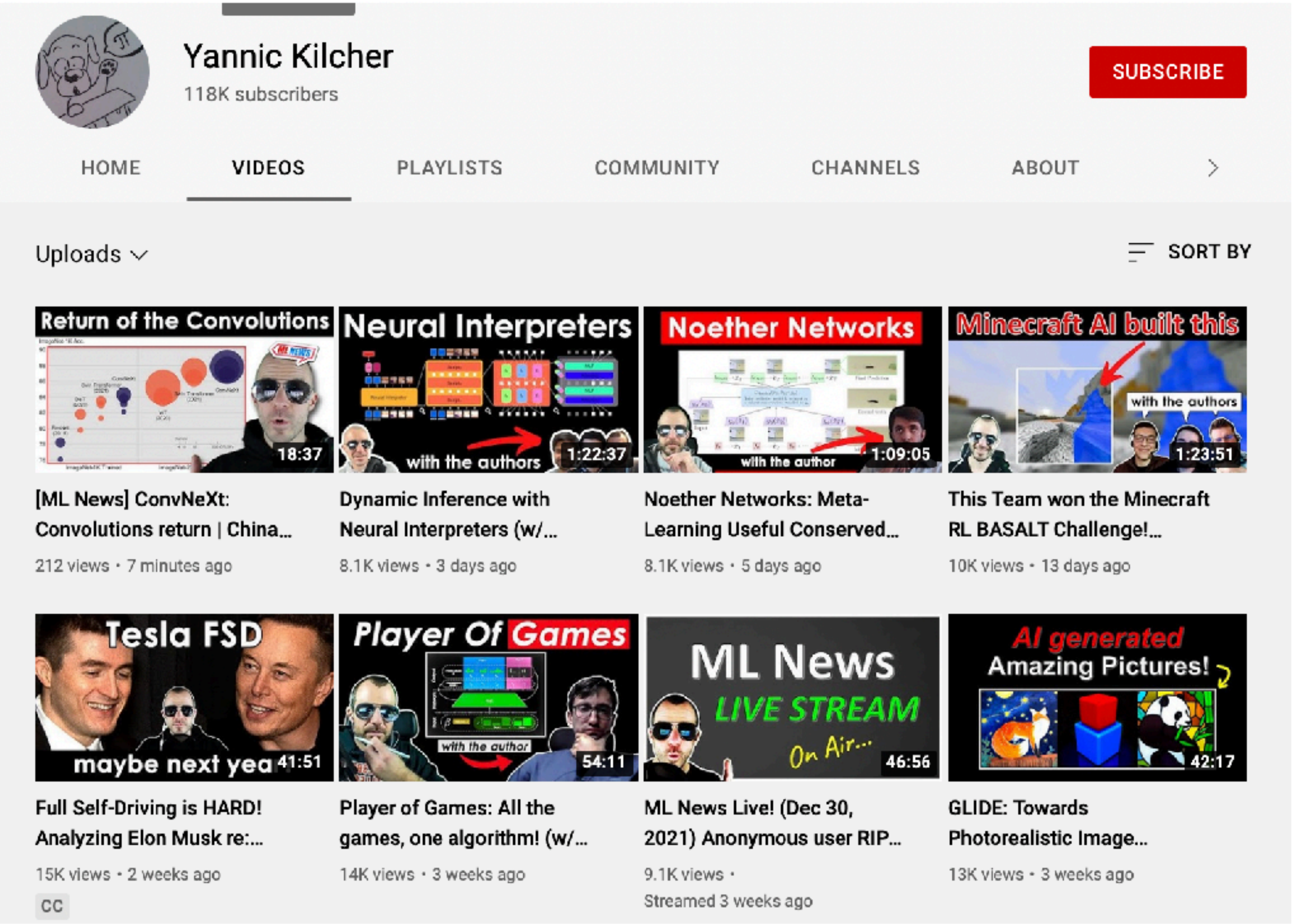


- Theory paper? Yes → sk
- Relevant for me as a use me. E.g. new non-linearity, optimizer, etc.
- Isn't the title over-keywo Adaptive DropBlock Enhanced Generative Ac Hyperspectral Image Classification". This kind erbalize.

Time for live example

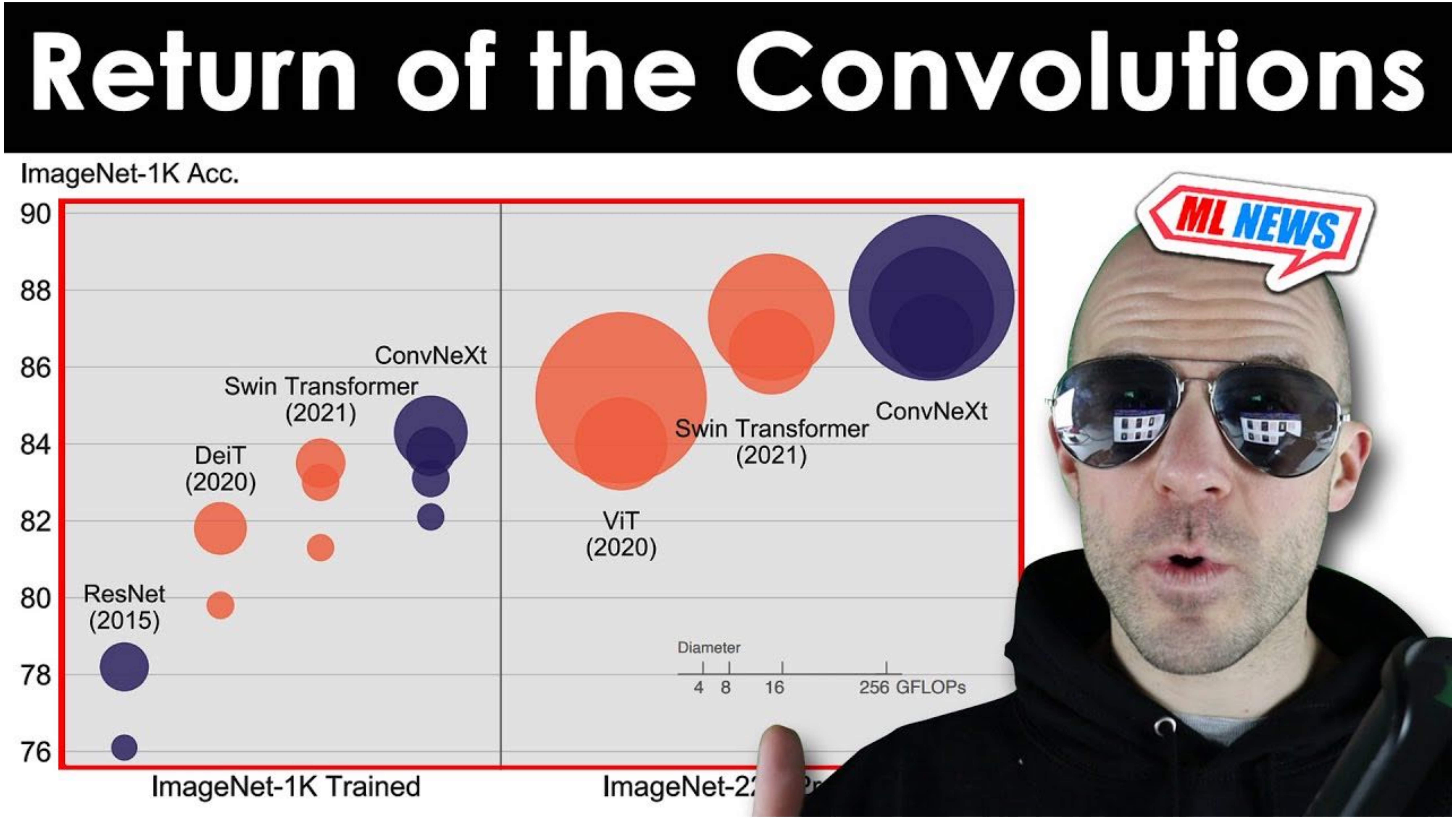
Additional paper sources

Yannic Kilcher



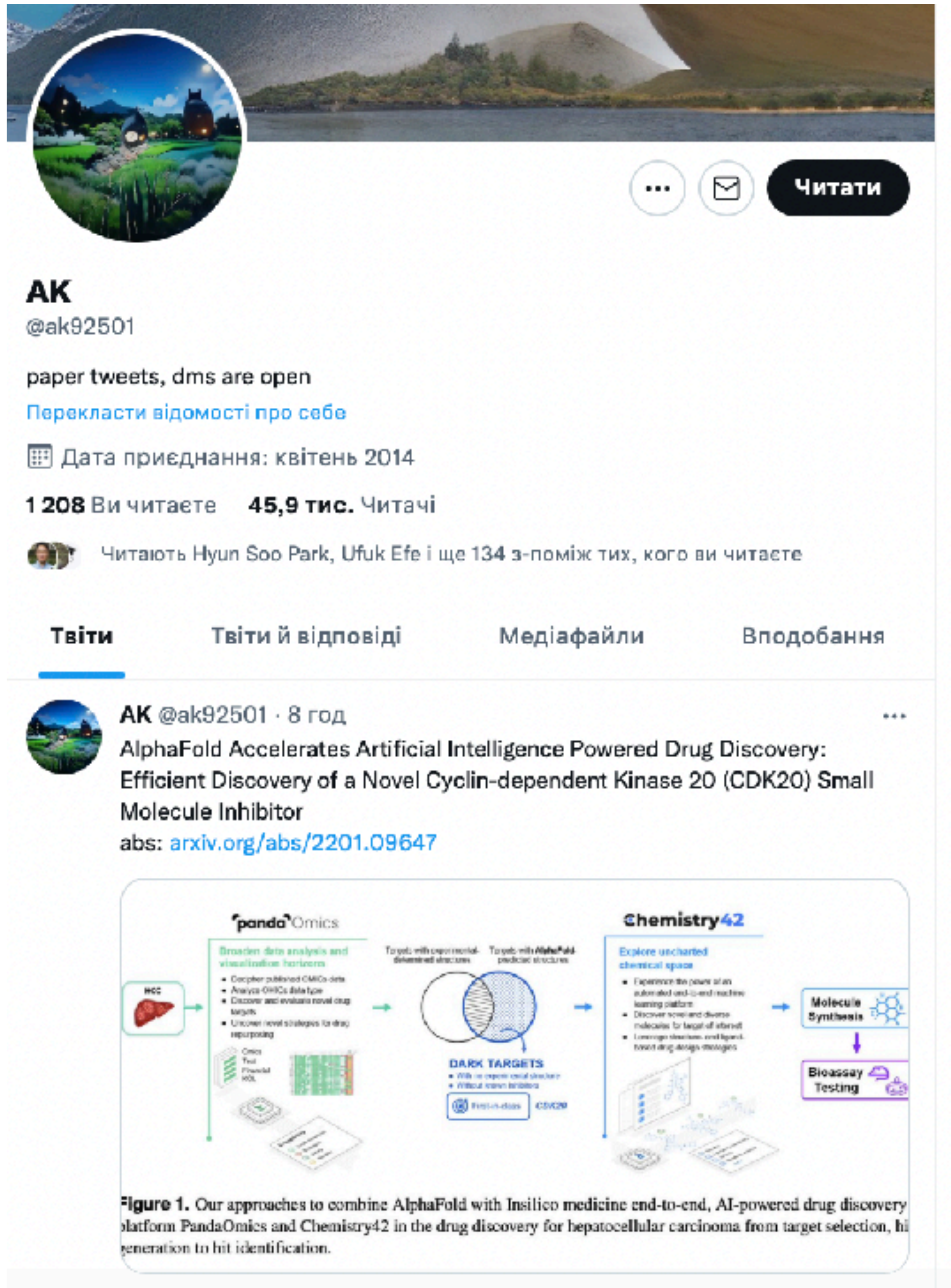
Additional paper sources

Yannic Kilcher



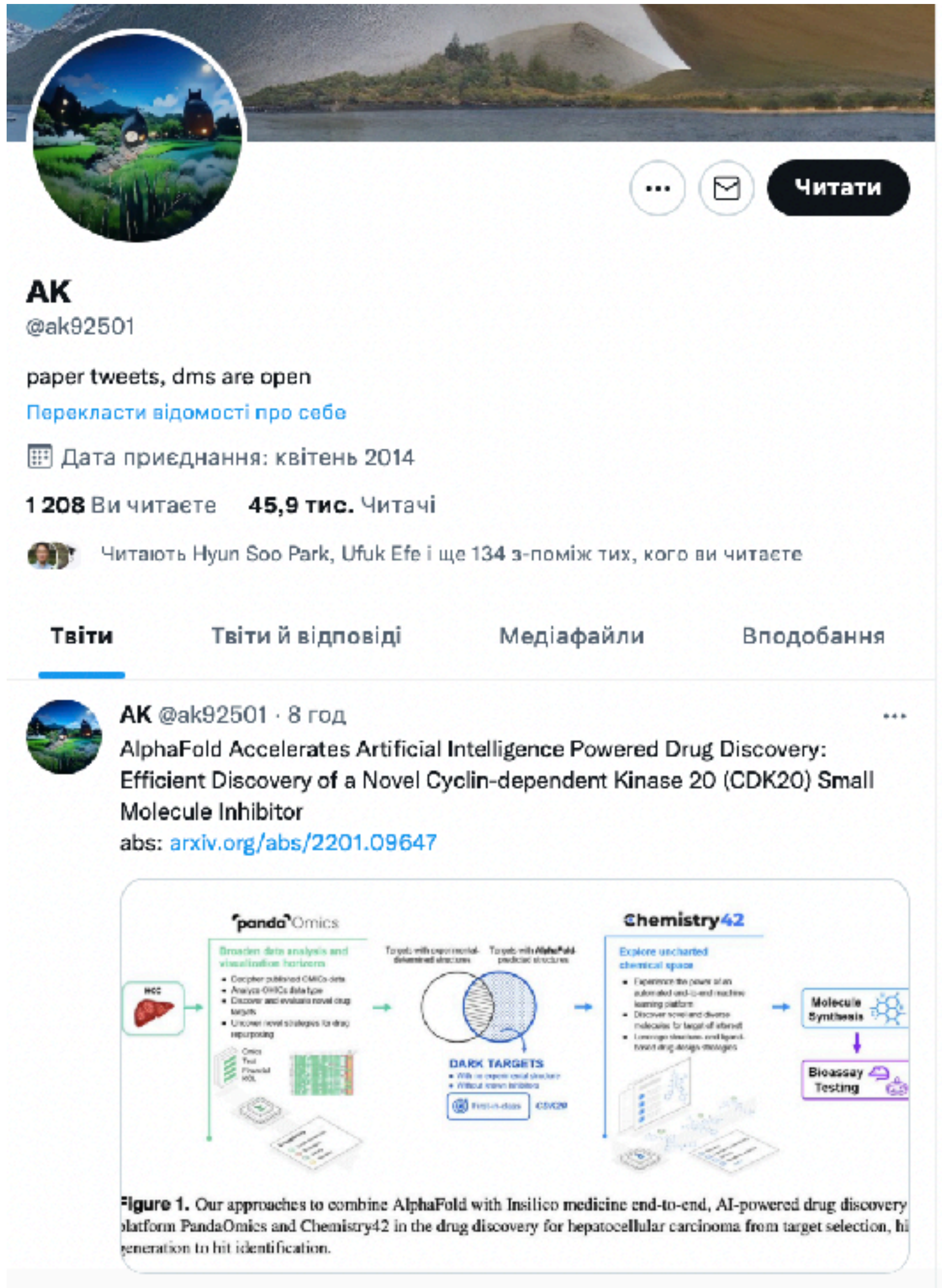
Twitter AK feed — general purpose interesting

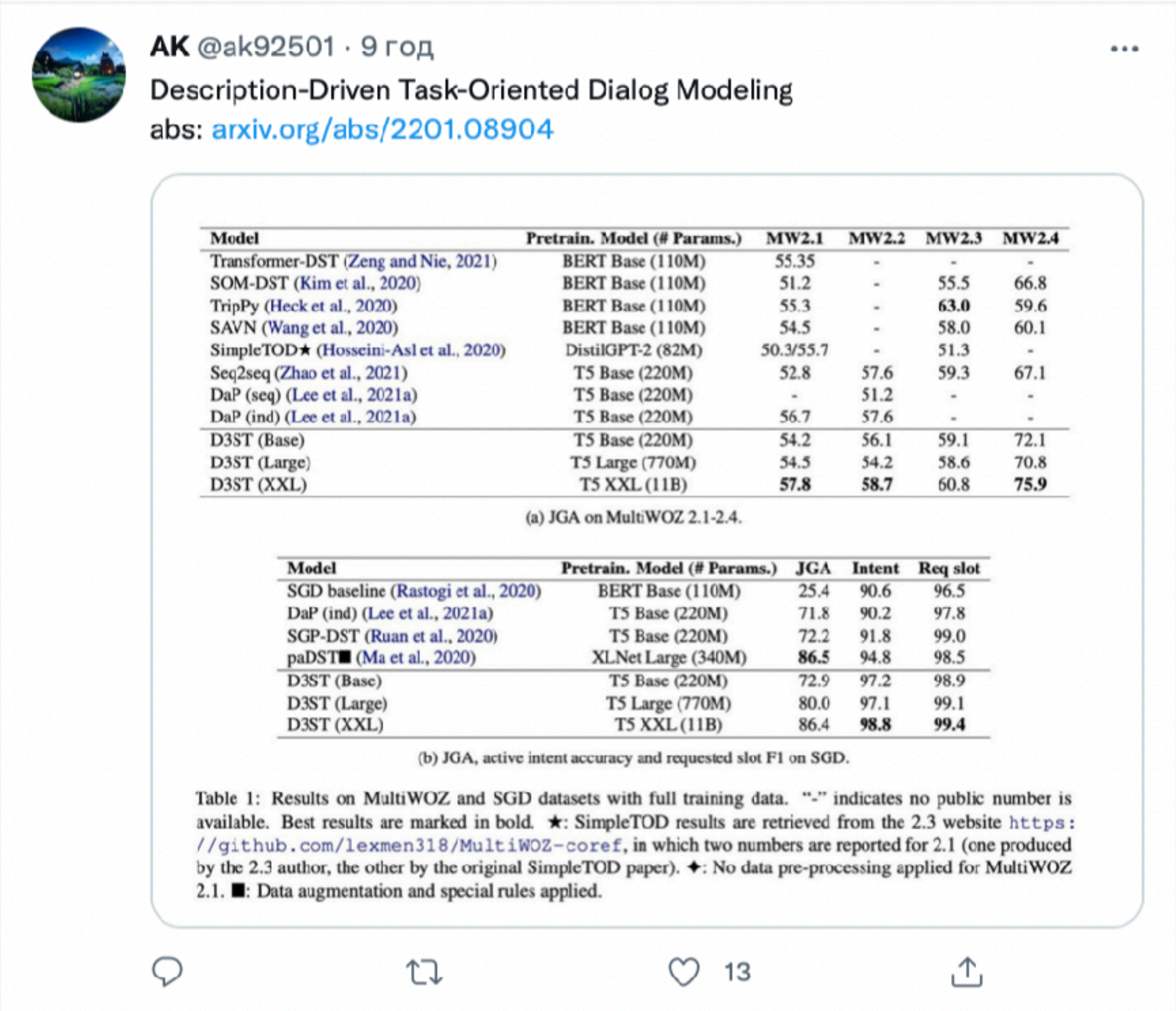
@ak92501



Twitter AK feed — general purpose interesting

@ak92501





Model	Pretrain. Model (# Params.)	MW2.1	MW2.2	MW2.3	MW2.4
Transformer-DST (Zeng and Nie, 2021)	BERT Base (110M)	55.35	-	-	-
SOM-DST (Kim et al., 2020)	BERT Base (110M)	51.2	-	55.5	66.8
TripPy (Heck et al., 2020)	BERT Base (110M)	55.3	-	**63.0**	59.6
SAVN (Wang et al., 2020)	BERT Base (110M)	54.5	-	58.0	60.1
SimpleTOD★ (Hosseini-Asl et al., 2020)	DistilGPT-2 (82M)	50.3/55.7	-	51.3	-
Seq2seq (Zhao et al., 2021)	T5 Base (220M)	52.8	57.6	59.3	67.1
DaP (seq) (Lee et al., 2021a)	T5 Base (220M)	-	51.2	-	-
DaP (ind) (Lee et al., 2021a)	T5 Base (220M)	56.7	57.6	-	-
D3ST (Base)	T5 Base (220M)	54.2	56.1	59.1	72.1
D3ST (Large)	T5 Large (770M)	54.5	54.2	58.6	70.8
D3ST (XXL)	T5 XXL (11B)	**57.8**	**58.7**	60.8	**75.9**

(a) JGA on MultiWOZ 2.1-2.4.

Model	Pretrain. Model (# Params.)	JGA	Intent	Req slot
SGD baseline (Rastogi et al., 2020)	BERT Base (110M)	25.4	90.6	96.5
DaP (ind) (Lee et al., 2021a)	T5 Base (220M)	71.8	90.2	97.8
SGP-DST (Ruan et al., 2020)	T5 Base (220M)	72.2	91.8	99.0
paDST■ (Ma et al., 2020)	XLNet Large (340M)	**86.5**	94.8	98.5
D3ST (Base)	T5 Base (220M)	72.9	97.2	98.9
D3ST (Large)	T5 Large (770M)	80.0	97.1	99.1
D3ST (XXL)	T5 XXL (11B)	86.4	**98.8**	**99.4**

(b) JGA, active intent accuracy and requested slot F1 on SGD.

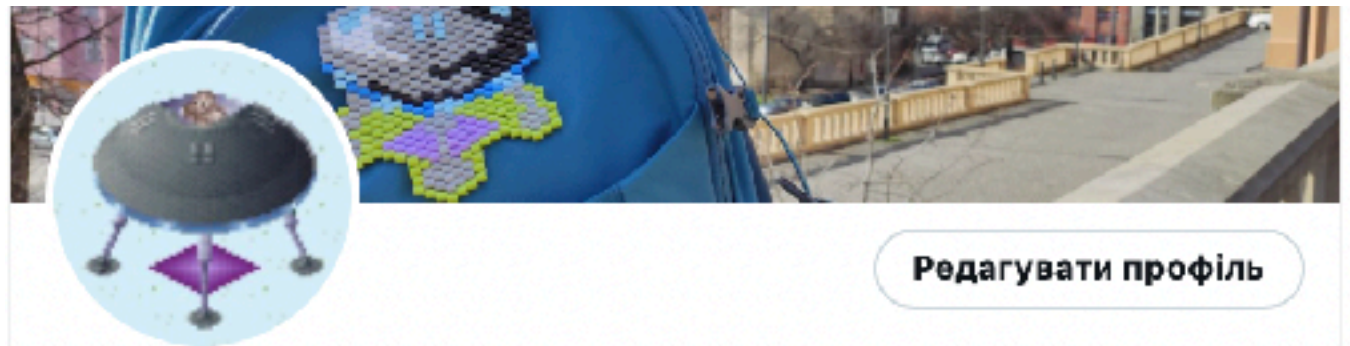
Table 1: Results on MultiWOZ and SGD datasets with full training data. "-" indicates no public number is available. Best results are marked in bold. ★: SimpleTOD results are retrieved from the 2.3 website `https://github.com/lexmen318/MultiWOZ-coref`, in which two numbers are reported for 2.1 (one produced by the 2.3 author, the other by the original SimpleTOD paper). ✦: No data pre-processing applied for MultiWOZ 2.1. ■: Data augmentation and special rules applied.

Twitter my feed: Image matching & hype

@ducha_aiki

Twitter my feed: Image matching & hype

@ducha_aiki



Редагувати профіль

Dmytro Mishkin
@ducha_aiki

Marrying classical CV and Deep Learning. I do things, which work, rather than being novel, but not working.
Перекласти відомості про себе

Ukraine dmytro.ai Дата приєднання: липень 2017

441 Ви читаєте **5 852** Читачі

Твіти Твіти й відповіді Медіафайли Вподобання

Закріплений твіт
Dmytro Mishkin @ducha_aiki · 13 жовт. 2020
ArXiving before submission :

- helps professional identity building;
-protects against idea re-discovery/theft/gate-keeping
- facilitates open research distribution
- reduces inequality.

our refined statement, with @amy_tabb & Jiri Matas, now on arXiv

arxiv.org
ArXiving Before Submission Helps Everyone
We claim, and present evidence, that allowing arXiv publication before a conference or journal ...

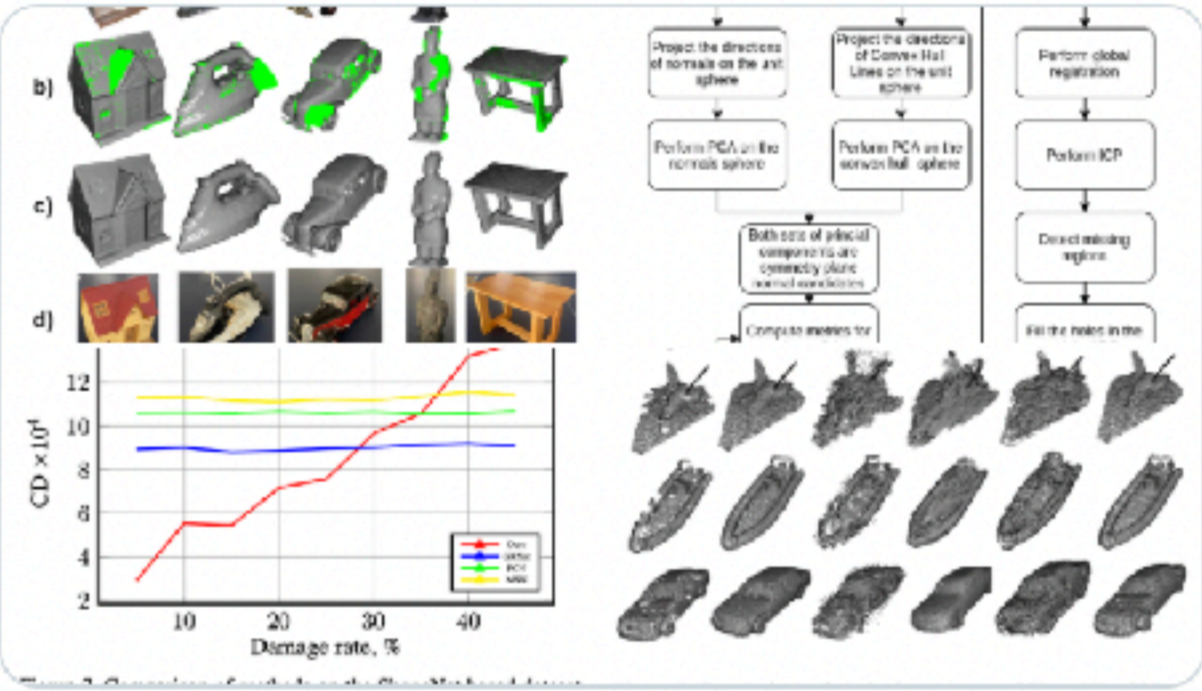
9 44 160

Dmytro Mishkin @ducha_aiki · 7 січ.
Towards realistic symmetry-based completion of previously unseen point clouds

Taras Rumezhak, @dobosevych , Rostyslav Hryniv, Vladyslav Selotkin, Volodymyr Karpiv, Mykola Maksymenko, @ucu_apps

tl;dr: good handcrafted method for point cloud completion

arxiv.org/abs/2201.01858



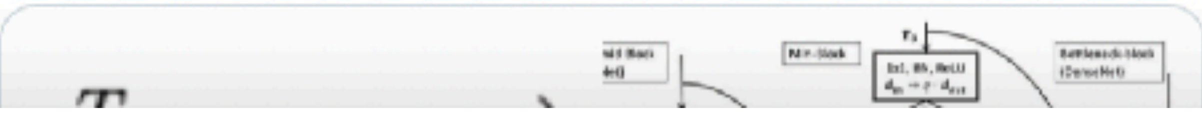
1 7 15

Dmytro Mishkin @ducha_aiki · 7 січ.
Bio-inspired Min-Nets Improve the Performance and Robustness of Deep Networks

Philipp Grüning, Erhardt Barth

tl;dr: out = min(ReLU(convA(BN(x))), ReLU(convB(BN((x)))
Idea: both filters should fire for reliable detection, kind of opposite of maxout idea.

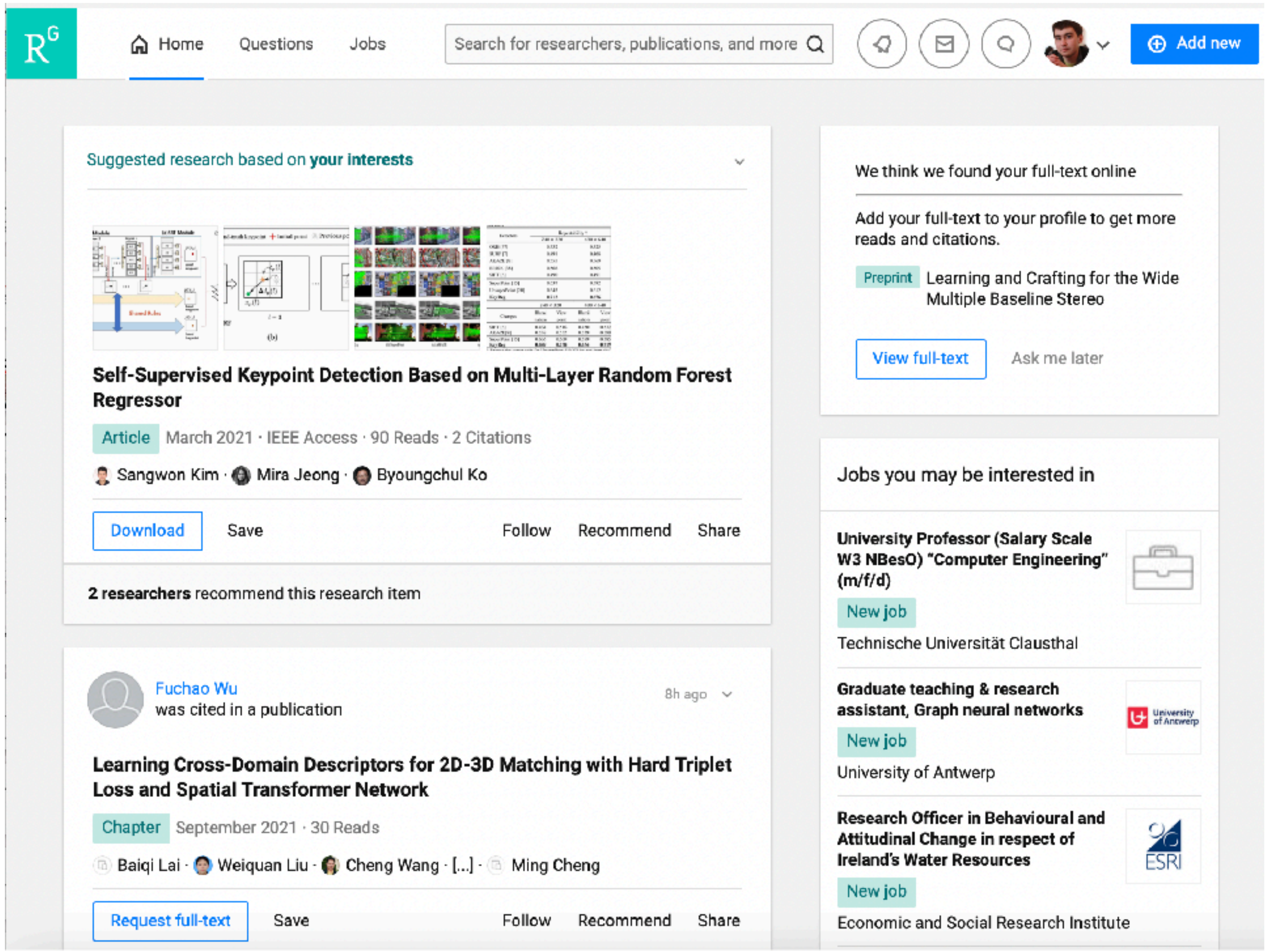
arxiv.org/abs/2201.02149...



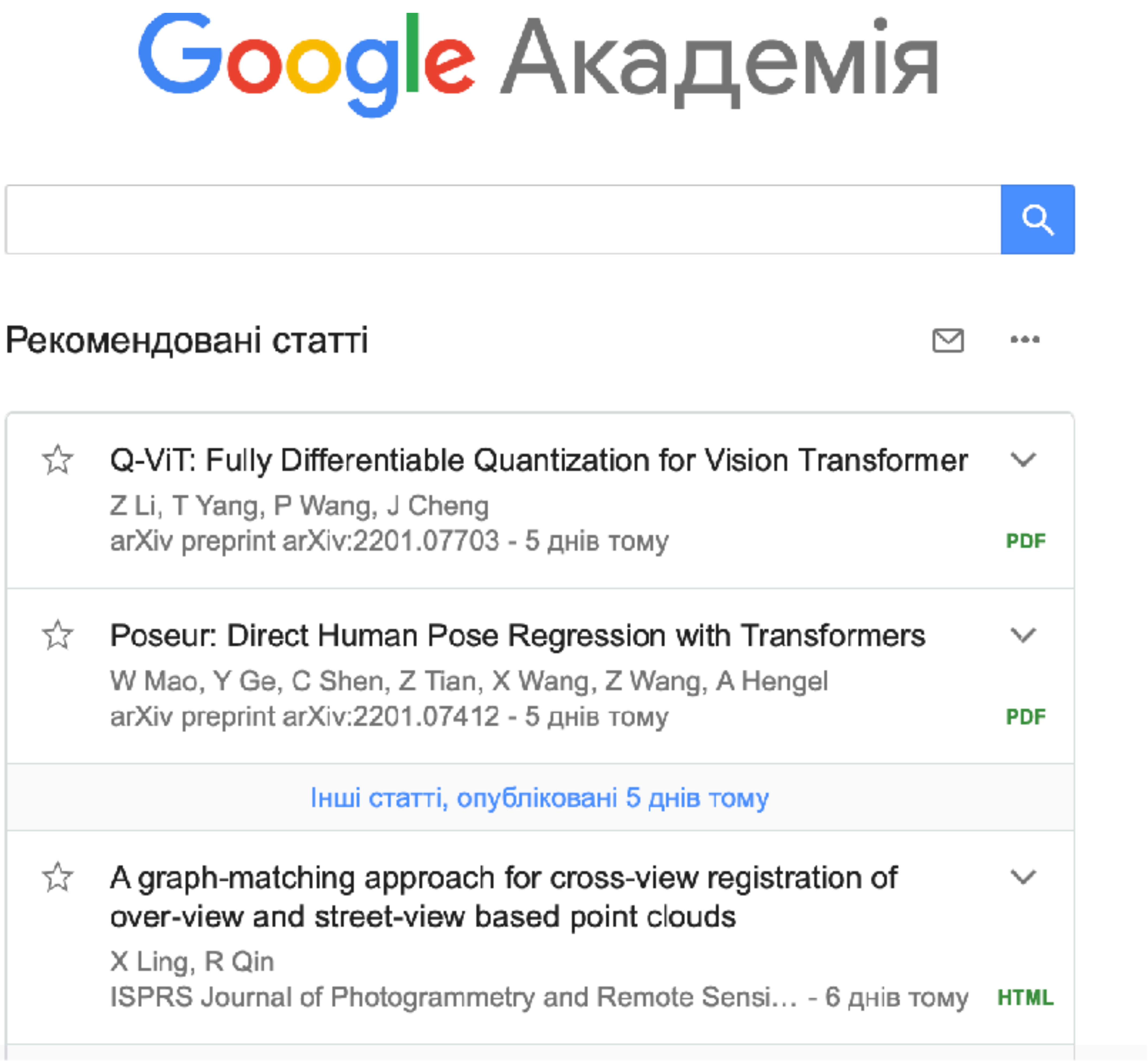
Register on twitter and follow relevant people

ResearchGate

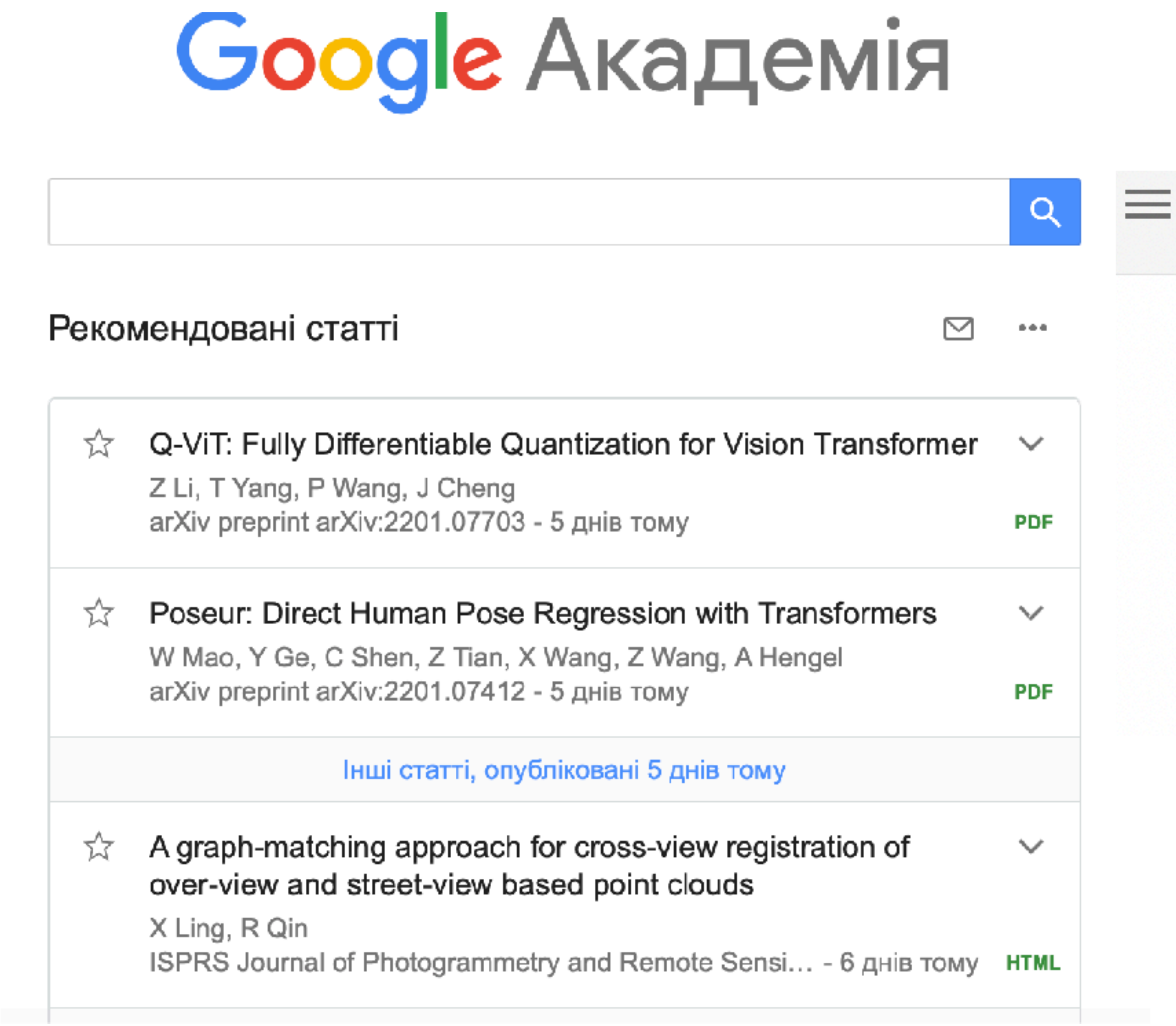
way to share and request paywalled papers

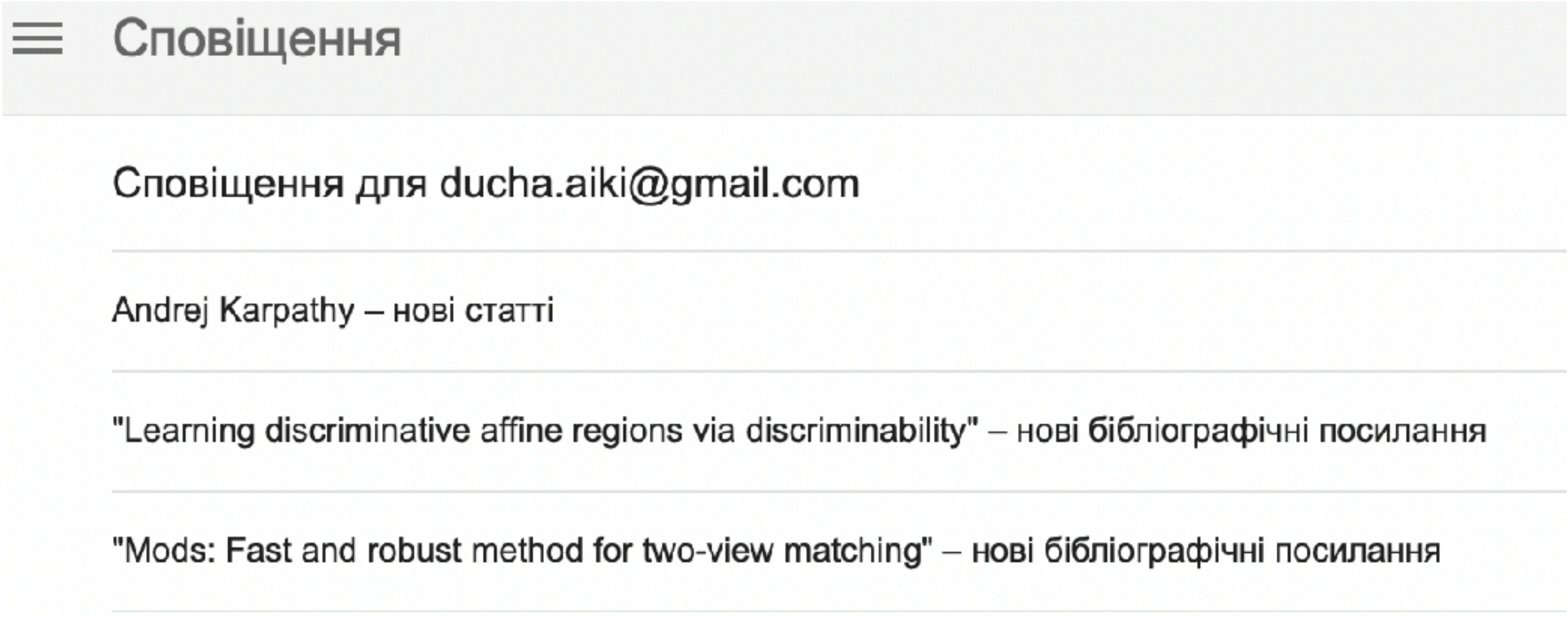


Google Scholar recommendations



Google Scholar recommendations





How to read a method paper?

The reading order of the method paper

Working hard to know your neighbor's margins: Local descriptor learning loss

Thanks to François Fleuret and Darshan Deshpande for help with this slide

The reading order of the method paper

Working hard to know your neighbor’s margins: Local descriptor learning loss

- Abstract

Thanks to François Fleuret and Darshan Deshpande for help with this slide

The reading order of the method paper

Working hard to know your neighbor's margins: Local descriptor learning loss

- Abstract

Abstract

We introduce a loss for metric learning, which is inspired by the Lowe's matching criterion for SIFT. We show that the proposed loss, that maximizes the distance between the closest positive and closest negative example in the batch, is better than complex regularization methods; it works well for both shallow and deep convolution network architectures. Applying the novel loss to the L2Net CNN architecture results in a compact descriptor named HardNet. It has the same dimensionality as SIFT (128) and shows state-of-art performance in wide baseline stereo, patch verification and instance retrieval benchmarks.

Thanks to François Fleuret and Darshan Deshpande for help with this slide

The reading order of the method paper

Working hard to know your neighbor's margins: Local descriptor learning loss

- Abstract
- Method figure

Abstract

We introduce a loss for metric learning, which is inspired by the Lowe's matching criterion for SIFT. We show that the proposed loss, that maximizes the distance between the closest positive and closest negative example in the batch, is better than complex regularization methods; it works well for both shallow and deep convolution network architectures. Applying the novel loss to the L2Net CNN architecture results in a compact descriptor named HardNet. It has the same dimensionality as SIFT (128) and shows state-of-art performance in wide baseline stereo, patch verification and instance retrieval benchmarks.

Thanks to François Fleuret and Darshan Deshpande for help with this slide

The reading order of the method paper

Working hard to know your neighbor's margins: Local descriptor learning loss

- Abstract
- Method figure

Abstract

We introduce a loss for metric learning, which is inspired by the Lowe's matching criterion for SIFT. We show that the proposed loss, that maximizes the distance between the closest positive and closest negative example in the batch, is better than complex regularization methods; it works well for both shallow and deep convolution network architectures. Applying the novel loss to the L2Net CNN architecture results in a compact descriptor named HardNet. It has the same dimensionality as SIFT (128) and shows state-of-art performance in wide baseline stereo, patch verification and instance retrieval benchmarks.

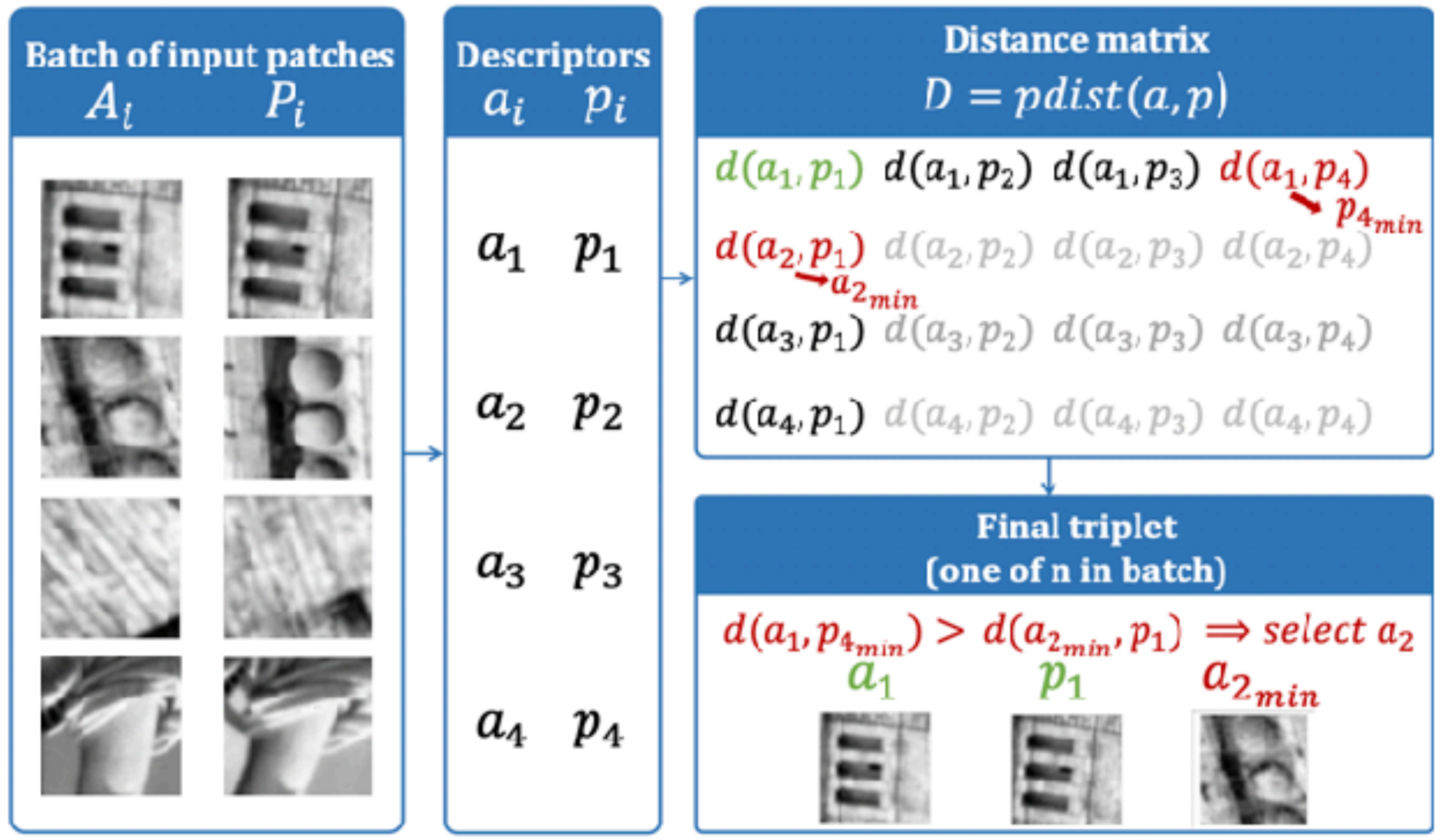


Figure 1: Proposed sampling procedure. First, patches are described by the current network, then a distance matrix is calculated. The closest non-matching descriptor – shown in red – is selected for each a_i and p_i patch from positive pair (green) respectively. Finally, among two negative candidates the hardest one is chosen. All operations are done in a single forward pass.

Thanks to François Fleuret and Darshan Deshpande for help with this slide

The reading order of the method paper

Working hard to know your neighbor's margins: Local descriptor learning loss

- Abstract
- Method figure

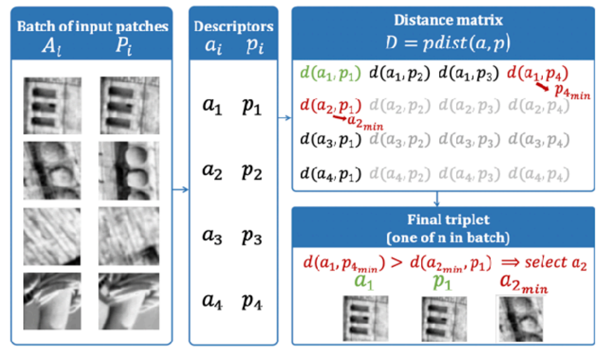


Figure 1: Proposed sampling procedure. First, patches are described by the current network, then a distance matrix is calculated. The closest non-matching descriptor – shown in red – is selected for each a_i and p_i patch from positive pair (green) respectively. Finally, among two negative candidates the hardest one is chosen. All operations are done in a single forward pass.

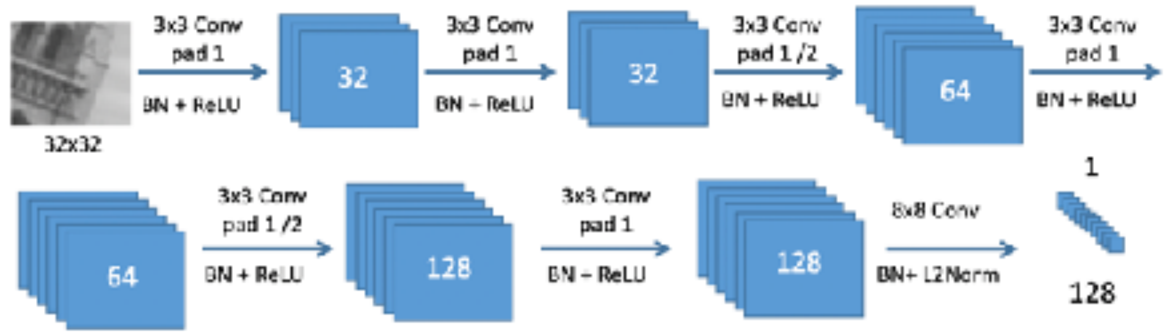


Figure 2: The architecture of our network, adopted from L2Net [24]. Each convolutional layer is followed by batch normalization and ReLU, except the last one. Dropout regularization is used before the last convolution layer.

Thanks to François Fleuret and Darshan Deshpande for help with this slide

The reading order of the method paper

Working hard to know your neighbor’s margins: Local descriptor learning loss

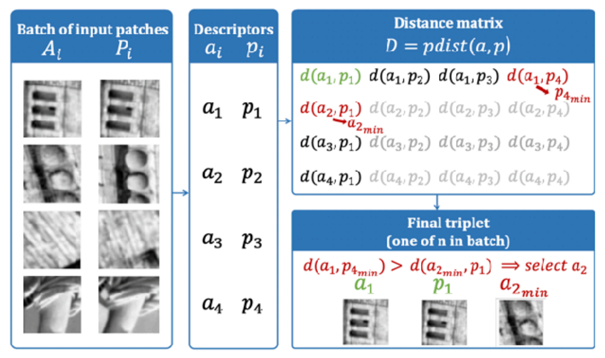


Figure 1: Proposed sampling procedure. First, patches are described by the current network, then a distance matrix is calculated. The closest non-matching descriptor – shown in red – is selected for each a_i and p_i patch from positive pair (green) respectively. Finally, among two negative candidates the hardest one is chosen. All operations are done in a single forward pass.

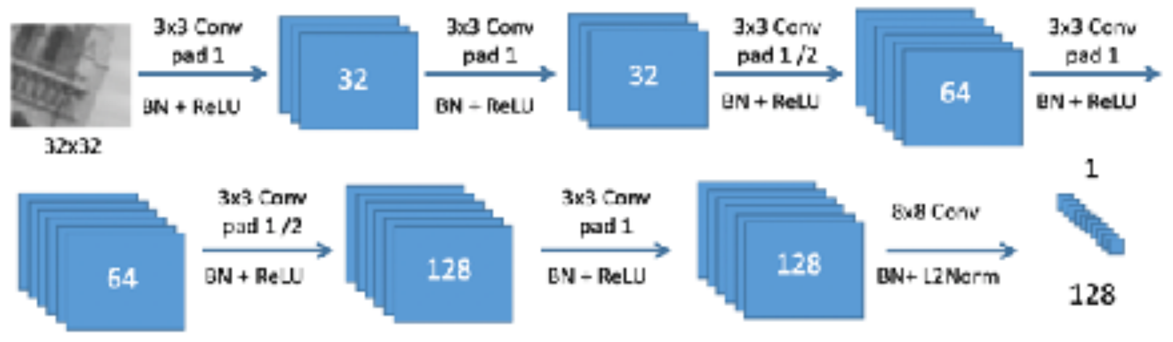


Figure 2: The architecture of our network, adopted from L2Net [24]. Each convolutional layer is followed by batch normalization and ReLU, except the last one. Dropout regularization is used before the last convolution layer.

- Abstract
- Method figure
- Main plot of results

Thanks to François Fleuret and Darshan Deshpande for help with this slide

The reading order of the method paper

Working hard to know your neighbor’s margins: Local descriptor learning loss

- Abstract
- Method figure
- Main plot of results

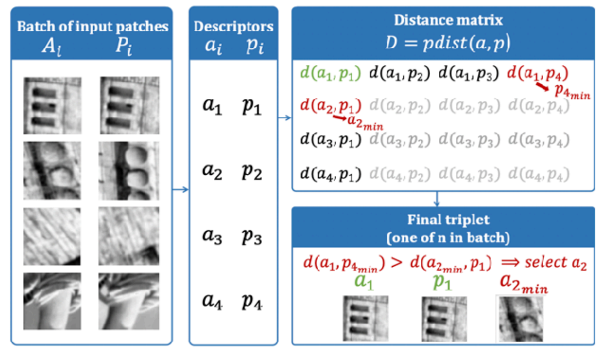


Figure 1: Proposed sampling procedure. First, patches are described by the current network, then a distance matrix is calculated. The closest non-matching descriptor – shown in red – is selected for each a_i and p_i patch from positive pair (green) respectively. Finally, among two negative candidates the hardest one is chosen. All operations are done in a single forward pass.

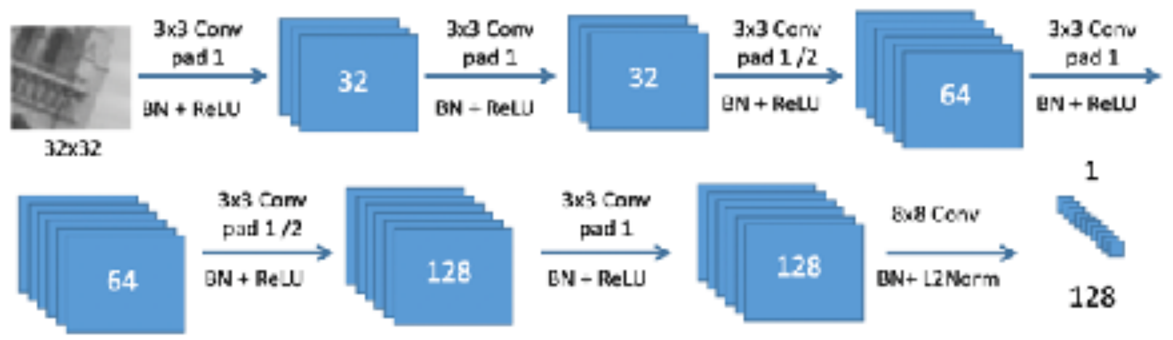
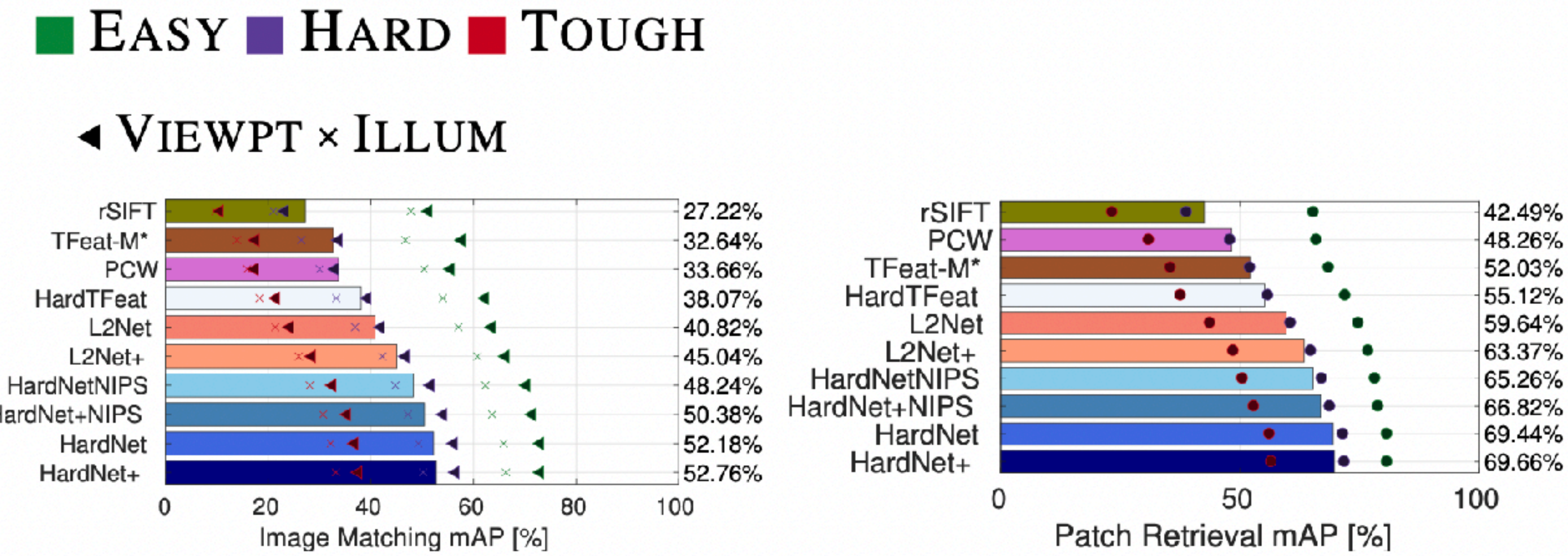


Figure 2: The architecture of our network, adopted from L2Net [24]. Each convolutional layer is followed by batch normalization and ReLU, except the last one. Dropout regularization is used before the last convolution layer.



Thanks to François Fleuret and Darshan Deshpande for help with this slide

The reading order of the method paper

Working hard to know your neighbor’s margins: Local descriptor learning loss

- Abstract
- Method figure
- Main plot of results
- Conclusion

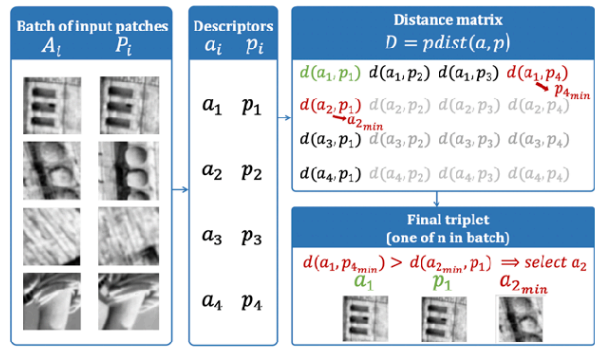


Figure 1: Proposed sampling procedure. First, patches are described by the current network, then a distance matrix is calculated. The closest non-matching descriptor – shown in red – is selected for each a_i and p_i patch from positive pair (green) respectively. Finally, among two negative candidates the hardest one is chosen. All operations are done in a single forward pass.

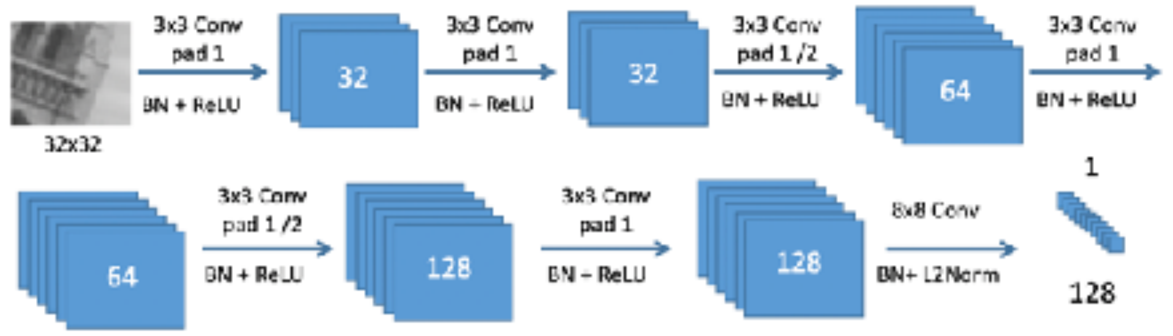
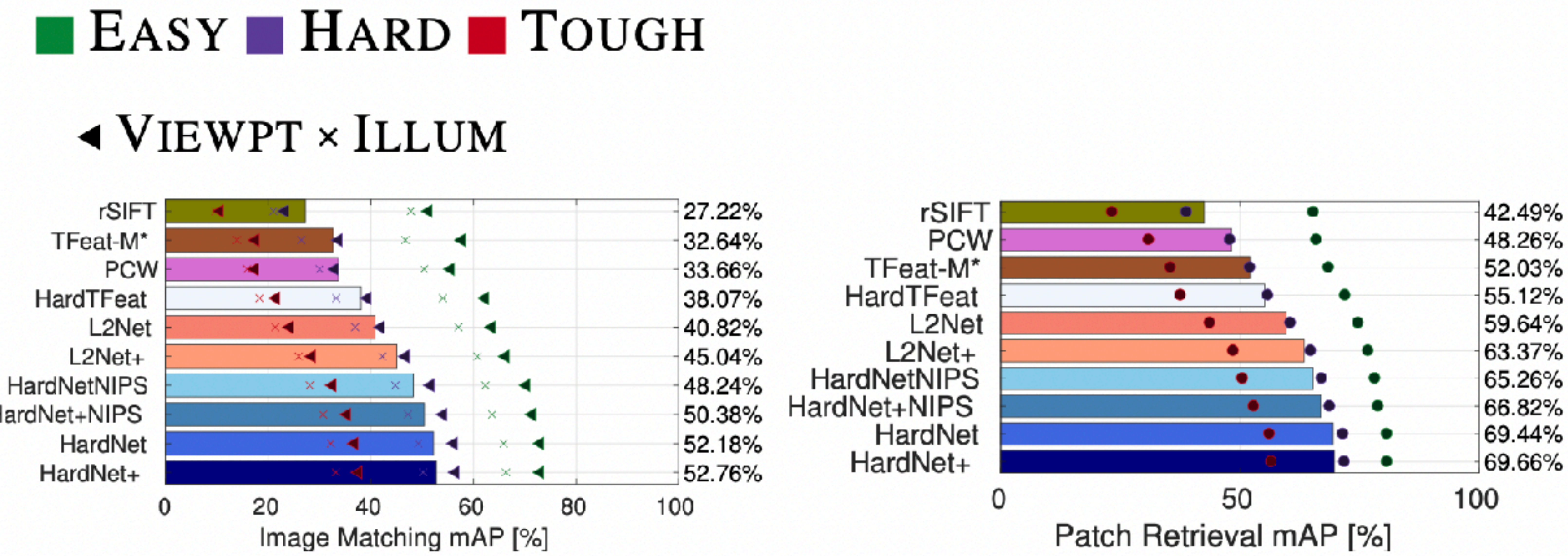


Figure 2: The architecture of our network, adopted from L2Net [24]. Each convolutional layer is followed by batch normalization and ReLU, except the last one. Dropout regularization is used before the last convolution layer.



Thanks to François Fleuret and Darshan Deshpande for help with this slide

The reading order of the method paper

Working hard to know your neighbor’s margins: Local descriptor learning loss

- Abstract
- Method figure
- Main plot of results
- Conclusion
  - If it is not rephrased abstract

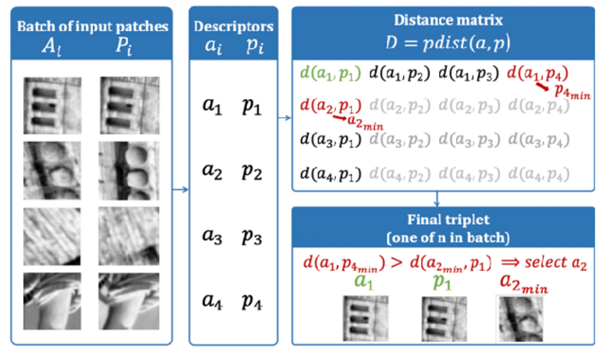


Figure 1: Proposed sampling procedure. First, patches are described by the current network, then a distance matrix is calculated. The closest non-matching descriptor – shown in red – is selected for each a_i and p_i patch from positive pair (green) respectively. Finally, among two negative candidates the hardest one is chosen. All operations are done in a single forward pass.

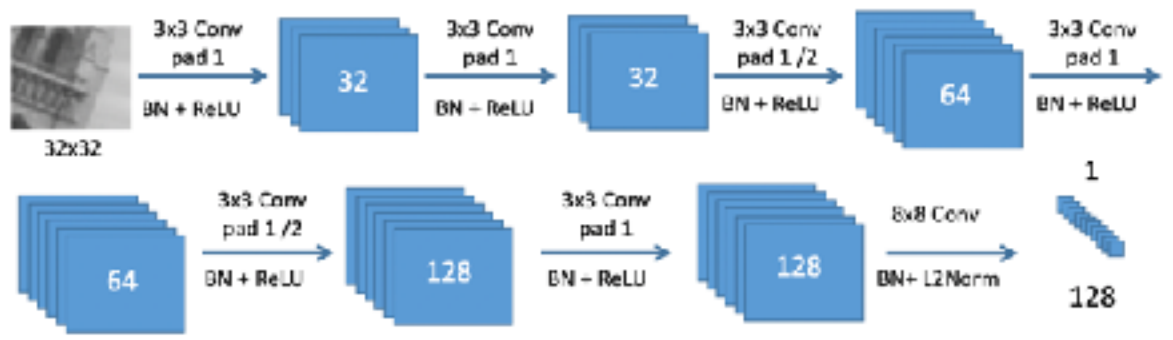
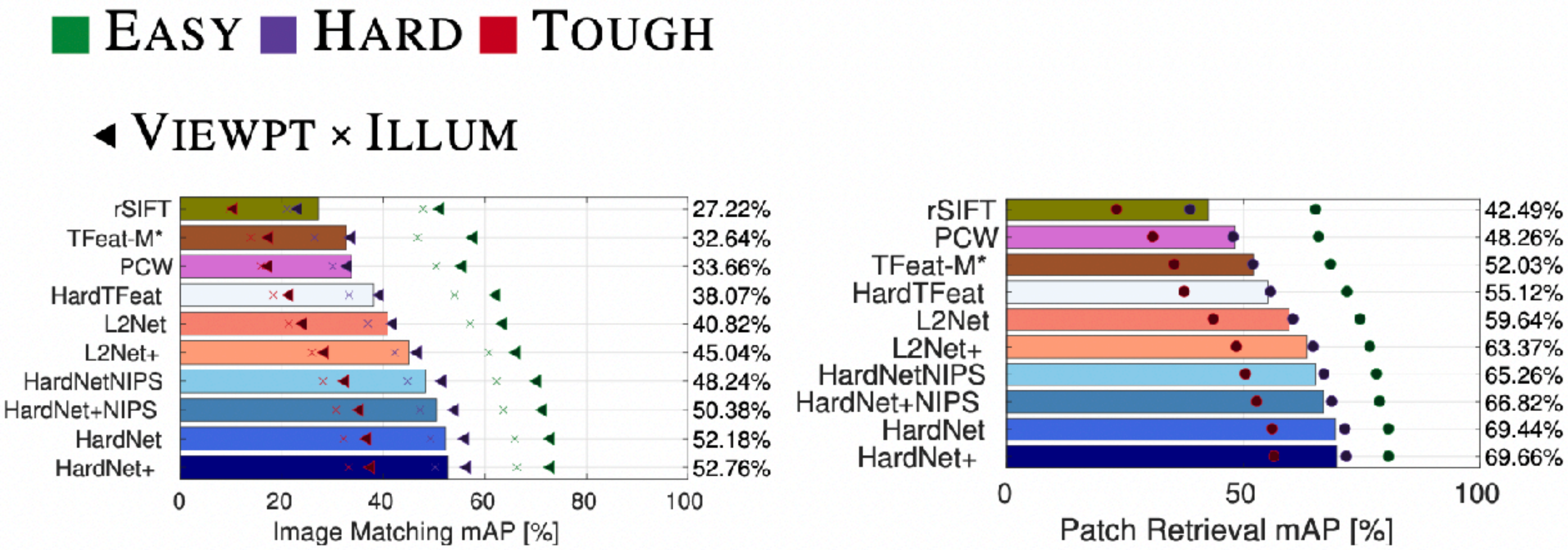


Figure 2: The architecture of our network, adopted from L2Net [24]. Each convolutional layer is followed by batch normalization and ReLU, except the last one. Dropout regularization is used before the last convolution layer.



Thanks to François Fleuret and Darshan Deshpande for help with this slide

The reading order of the method paper

Working hard to know your neighbor’s margins: Local descriptor learning loss

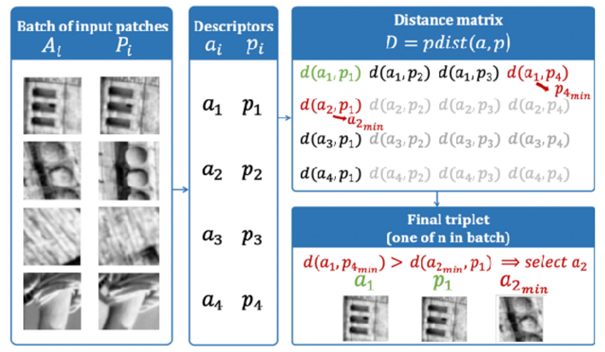


Figure 1: Proposed sampling procedure. First, patches are described by the current network, then a distance matrix is calculated. The closest non-matching descriptor – shown in red – is selected for each a_i and p_i patch from positive pair (green) respectively. Finally, among two negative candidates the hardest one is chosen. All operations are done in a single forward pass.

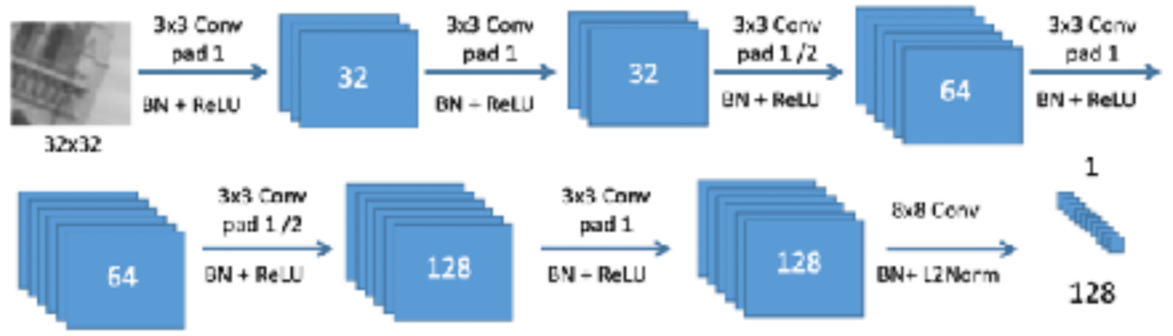
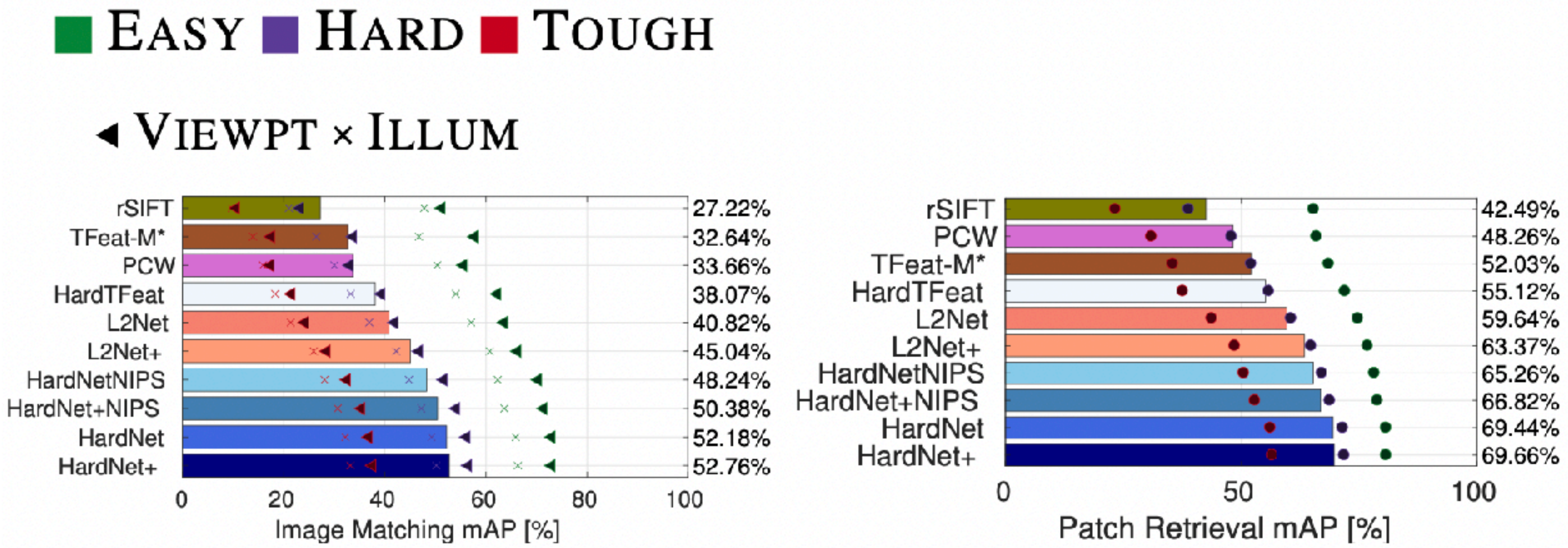


Figure 2: The architecture of our network, adopted from L2Net [24]. Each convolutional layer is followed by batch normalization and ReLU, except the last one. Dropout regularization is used before the last convolution layer.

- Abstract
- Method figure
- Main plot of results
- Conclusion
  - If it is not rephrased abstract
- Related work



Thanks to François Fleuret and Darshan Deshpande for help with this slide

Advice by Dushyant Mehta

From personal painful experience

Source: https://twitter.com/mehtadushy/status/1482424896961536007

Advice by Dushyant Mehta

From personal painful experience

1. Don't trust the story the paper tries to sell.

Source: https://twitter.com/mehtadushy/status/1482424896961536007

Advice by Dushyant Mehta

From personal painful experience

1. Don't trust the story the paper tries to sell.

   - supposed intuition is often not true

Source: https://twitter.com/mehtadushy/status/1482424896961536007

Advice by Dushyant Mehta

From personal painful experience

1. Don't trust the story the paper tries to sell.
    - supposed intuition is often not true
    - Ask instead, what could be the other reasons for the claimed improvements.

Source: https://twitter.com/mehtadushy/status/1482424896961536007

Advice by Dushyant Mehta

From personal painful experience

1. Don't trust the story the paper tries to sell.
    - supposed intuition is often not true
    - Ask instead, what could be the other reasons for the claimed improvements.
2. Learn to ask yourself about simple and obvious baselines.

Source: https://twitter.com/mehtadushy/status/1482424896961536007

Advice by Dushyant Mehta

From personal painful experience

1. Don't trust the story the paper tries to sell.
    - supposed intuition is often not true
    - Ask instead, what could be the other reasons for the claimed improvements.
2. Learn to ask yourself about simple and obvious baselines.
    - What should have been the simple and obvious baselines for the paper you are reading. How would you solve the task building from first principles?

Source: https://twitter.com/mehtadushy/status/1482424896961536007

Advice by Dushyant Mehta

From personal painful experience

1. Don't trust the story the paper tries to sell.
    - supposed intuition is often not true
    - Ask instead, what could be the other reasons for the claimed improvements.
2. Learn to ask yourself about simple and obvious baselines.
    - What should have been the simple and obvious baselines for the paper you are reading. How would you solve the task building from first principles?
    - What experiments would you design to validate the claims of the paper?

Source: https://twitter.com/mehtadushy/status/1482424896961536007

Advice by Dushyant Mehta

From personal painful experience

Source: https://twitter.com/mehtadushy/status/1482424896961536007

Advice by Dushyant Mehta

From personal painful experience

3. Don't take the results at their face value

Source: https://twitter.com/mehtadushy/status/1482424896961536007

Advice by Dushyant Mehta

From personal painful experience

3. Don't take the results at their face value

   - Sure there is a big convenient table with all manner of methods conveniently compared.

Source: https://twitter.com/mehtadushy/status/1482424896961536007

Advice by Dushyant Mehta

From personal painful experience

3. Don't take the results at their face value

    - Sure there is a big convenient table with all manner of methods conveniently compared.

    - Are they really comparable? Same data, same compute budget, fair training, fair evaluation?

Source: https://twitter.com/mehtadushy/status/1482424896961536007

Advice by Dushyant Mehta

From personal painful experience

3. Don't take the results at their face value

    - Sure there is a big convenient table with all manner of methods conveniently compared.

    - Are they really comparable? Same data, same compute budget, fair training, fair evaluation?

4. Don't discount the simple methods even if they are slightly worse than overly complex (claimed) SoTA.

Source: https://twitter.com/mehtadushy/status/1482424896961536007

Advice by Dushyant Mehta

From personal painful experience

3. Don't take the results at their face value

    - Sure there is a big convenient table with all manner of methods conveniently compared.

    - Are they really comparable? Same data, same compute budget, fair training, fair evaluation?

4. Don't discount the simple methods even if they are slightly worse than overly complex (claimed) SoTA.

    - There is immense advantage to the simple. Less finicky training, less faff in deployment, faster experiment iteration,

Source: https://twitter.com/mehtadushy/status/1482424896961536007

Advice by Dushyant Mehta

From personal painful experience

Source: https://twitter.com/mehtadushy/status/1482424896961536007

Advice by Dushyant Mehta

From personal painful experience

5. Do NOT start building on top of their *method*

Source: https://twitter.com/mehtadushy/status/1482424896961536007

Advice by Dushyant Mehta

From personal painful experience

5. Do NOT start building on top of their *method*
   - At least without having determined to your satisfaction that the method is sound

Source: https://twitter.com/mehtadushy/status/1482424896961536007

Advice by Dushyant Mehta

From personal painful experience

5. Do NOT start building on top of their *method*
    - At least without having determined to your satisfaction that the method is sound
    - Reuse various parts from the codebase, without building on top of the method!

Source: https://twitter.com/mehtadushy/status/1482424896961536007

Advice by Dushyant Mehta

From personal painful experience

5. Do NOT start building on top of their *method*
    - At least without having determined to your satisfaction that the method is sound
    - Reuse various parts from the codebase, without building on top of the method!
6. Data, data, data! For a lot of the problems, not a lot beats good data.

Source: https://twitter.com/mehtadushy/status/1482424896961536007

Advice by Dushyant Mehta

From personal painful experience

5. Do NOT start building on top of their *method*
   - At least without having determined to your satisfaction that the method is sound
   - Reuse various parts from the codebase, without building on top of the method!
6. Data, data, data! For a lot of the problems, not a lot beats good data.
   - Not mindlessly collecting more data, but knowing the distribution of your data, and sampling smartly from it in a way that reflects your deployment setting.

Source: https://twitter.com/mehtadushy/status/1482424896961536007

Advice by Dushyant Mehta

From personal painful experience

5. Do NOT start building on top of their *method*
    - At least without having determined to your satisfaction that the method is sound
    - Reuse various parts from the codebase, without building on top of the method!
6. Data, data, data! For a lot of the problems, not a lot beats good data.
    - Not mindlessly collecting more data, but knowing the distribution of your data, and sampling smartly from it in a way that reflects your deployment setting.
    - Very simple/subtle sampling differences can cause a huge difference in your metrics, without anything about the method changing.

Source: https://twitter.com/mehtadushy/status/1482424896961536007

Advice by Dushyant Mehta

From personal painful experience

Source: https://twitter.com/mehtadushy/status/1482424896961536007

Advice by Dushyant Mehta

From personal painful experience

7. That brings us to the f**king metrics.

Source: https://twitter.com/mehtadushy/status/1482424896961536007

Advice by Dushyant Mehta

From personal painful experience

7. That brings us to the f**king metrics.

    - If you mindlessly optimize for one number, look at leaderboards based on one number, you haven't <u>understood the task</u>.

Source: https://twitter.com/mehtadushy/status/1482424896961536007

Advice by Dushyant Mehta

From personal painful experience

7. That brings us to the f**king metrics.

   - If you mindlessly optimize for one number, look at leaderboards based on one number, you haven't <u>understood the task</u>.

   - What does that one number tell you?

Source: https://twitter.com/mehtadushy/status/1482424896961536007

Advice by Dushyant Mehta

From personal painful experience

7. That brings us to the f**king metrics.
   - If you mindlessly optimize for one number, look at leaderboards based on one number, you haven't <u>understood the task</u>.
   - What does that one number tell you?
   - More importantly, what does it hide from you?

Source: https://twitter.com/mehtadushy/status/1482424896961536007

Advice by Dushyant Mehta

From personal painful experience

7. That brings us to the f**king metrics.
   - If you mindlessly optimize for one number, look at leaderboards based on one number, you haven't understood the task.
   - What does that one number tell you?
   - More importantly, what does it hide from you?
   - How relevant is it to the deployment scenarios for that task?

Source: https://twitter.com/mehtadushy/status/1482424896961536007

Advice by Dushyant Mehta

From personal painful experience

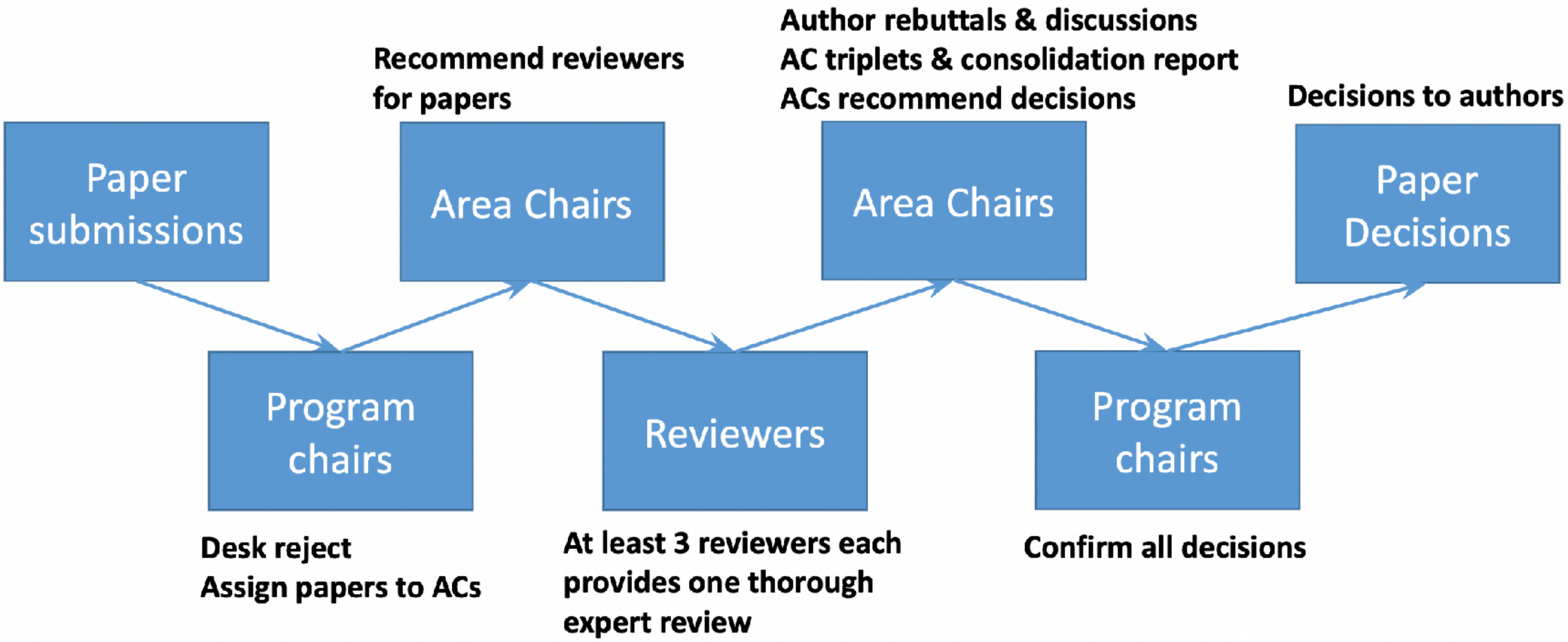
7. That brings us to the f**king metrics.

   - If you mindlessly optimize for one number, look at leaderboards based on one number, you haven't <u>understood the task</u>.
   - What does that one number tell you?
   - More importantly, what does it hide from you?
   - How relevant is it to the deployment scenarios for that task?
   - What metrics would better reflect what matters the most for your deployment scenario?

Source: https://twitter.com/mehtadushy/status/1482424896961536007

What does peer review mean?

The process of accepting/rejecting a paper



The role of reviewers in the process

- Provide an independent, objective, critical, and comprehensive review
  - **Key**: What is the knowledge advancement in the paper?
- Discuss with AC and reviewer buddies to (hopefully) reach consensus
  - Explain clearly the basis of your review and recommendation
  - It is OK if the reviewers disagree with one another even after discussions
  - AC will form recommendations weighing in reviews, rebuttals, and discussions
- Make your final recommendations with solid justifications
  - Read the rebuttal and discussions. Do they change your position? Why?
  - This facilitates the ACs to make final recommendations for the paper

What paper should be accepted?

- Any paper that, in accordance with CVPR community standards,
  - presents **sufficient knowledge advancement** that is **well grounded**;
  - is of **sufficient interest** to some **CVPR audiences** who could **benefit** from it

- Note: CVPR is very inclusive
  - Historically rejection solely for out-of-scope is rather rare

What should be included in the review?

- A concise summary of the paper
  - What problem is addressed in the paper?
  - Is it a new problem? If so, why does it matter? If not, why does it still matter?
  - What is the key to the solution? What is the main contribution?
  - Do the experiments sufficiently support the claims?
- A clear statement of strengths and weaknesses
  - What are the key contributions and why do they matter?
  - What aspects of the paper most need improvement?
- A comprehensive check of potential fundamental flaws in the paper
  - Are the assumptions and theories (mathematically) sound?
  - Are the experiments scientifically sound and valid?
  - Is the problem addressed trivial?
  - Did the paper miss important prior work? Has it been done before? If yes, where?

CVPR 2022 Reviewer tutorial

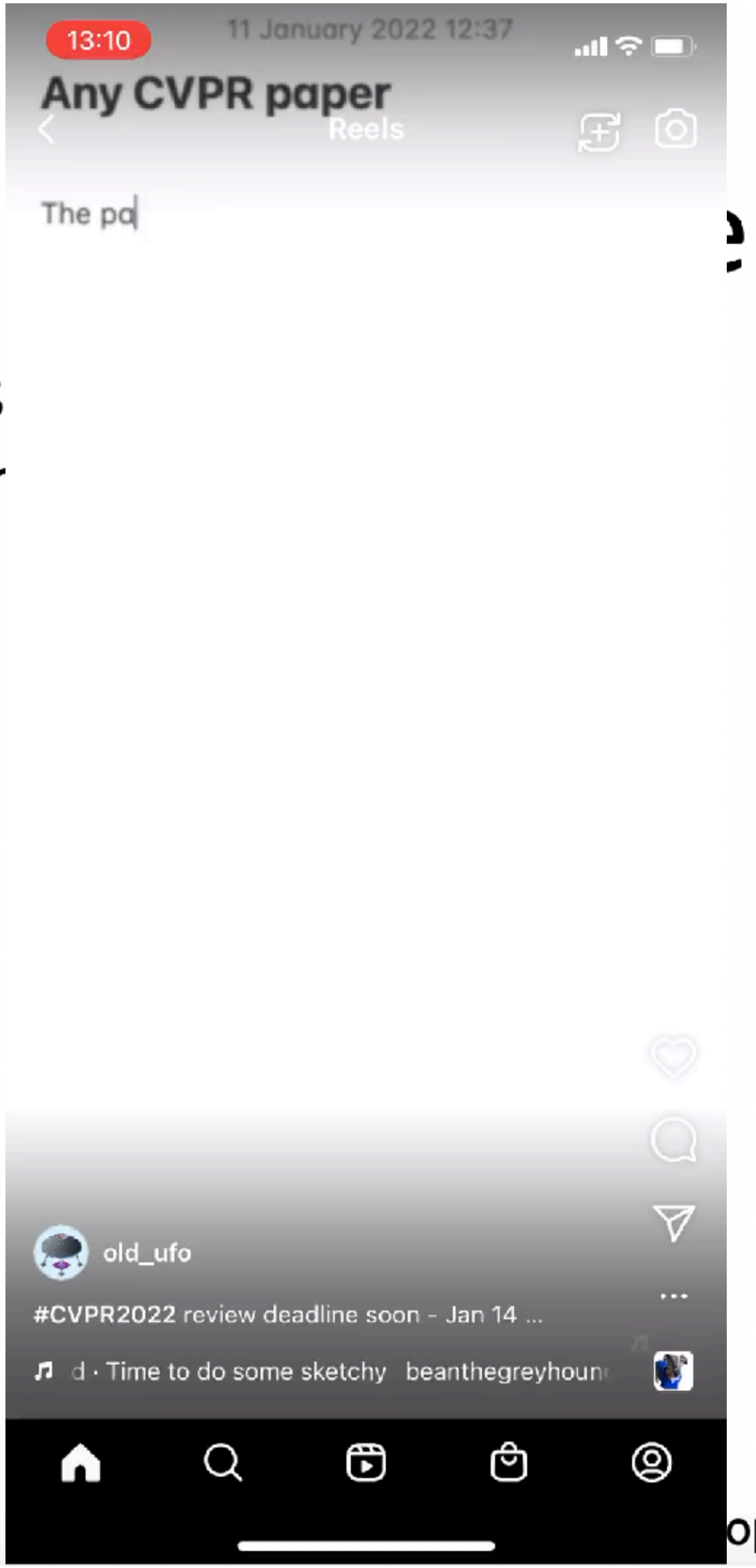
What should be avoided in the review?

- Common mistakes in a sloppy review
  - Arrogance, ignorance, and inaccuracy
    - Be responsible!
  - Pure opinions
    - Be grounded!
  - Novelty fallacy
    - Be knowledge-driven!
  - Blank assertions
    - Be substantial!
  - Policy entrepreneurism
    - Be sensible!
  - Intellectual laziness
    - Be active!

[Developed from: Reviewer Slides for CVPR’21]

What should be … e review?

- Common mistakes in a s
  - Arrogance, ignorance, ar
    - Be responsible!
  - Pure opinions
    - Be grounded!
  - Novelty fallacy
    - Be knowledge-driven!
  - Blank assertions
    - Be substantial!
  - Policy entrepreneurism
    - Be sensible!
  - Intellectual laziness
    - Be active!



oped from: Reviewer Slides for CVPR'21]

Functions of the conferences

Why authors submit

Functions of the conferences

Why authors submit

- **Knowledge dissemination**: "I want people to know about the new knowledge I discovered." The conference promises a certain minimum level of attention one's work gets. In other words, if you don't publish at top conferences, nobody would read your work.

https://ducha-aiki.medium.com/what-does-it-mean-to-publish-your-scientific-paper-in-2020-9ec3a8a90fe8 https://amytabb.com

Functions of the conferences
Why authors submit

- **Knowledge dissemination**: “I want people to know about the new knowledge I discovered.” The conference promises a certain minimum level of attention one’s work gets. In other words, if you don’t publish at top conferences, nobody would read your work.
- **Feedback**: I would like people to check my results through the review process and discussion.

https://ducha-aiki.medium.com/what-does-it-mean-to-publish-your-scientific-paper-in-2020-9ec3a8a90fe8 https://amytabb.com

Functions of the conferences

Why authors submit

- **Knowledge dissemination**: “I want people to know about the new knowledge I discovered.” The conference promises a certain minimum level of attention one’s work gets. In other words, if you don’t publish at top conferences, nobody would read your work.

- **Feedback**: I would like people to check my results through the review process and discussion.

- **Formal goodies**: checking boxes, required to defend Ph.D., for tenure package, performance review, to put into grant report, etc.

https://ducha-aiki.medium.com/what-does-it-mean-to-publish-your-scientific-paper-in-2020-9ec3a8a90fe8 https://amytabb.com

Functions of the conferences

Why authors submit

- **Knowledge dissemination**: “I want people to know about the new knowledge I discovered.” The conference promises a certain minimum level of attention one’s work gets. In other words, if you don’t publish at top conferences, nobody would read your work.
- **Feedback**: I would like people to check my results through the review process and discussion.
- **Formal goodies**: checking boxes, required to defend Ph.D., for tenure package, performance review, to put into grant report, etc.
- **Certification**. Publishing at CVPR is hard, therefore valuable.

https://ducha-aiki.medium.com/what-does-it-mean-to-publish-your-scientific-paper-in-2020-9ec3a8a90fe8 https://amytabb.com

Functions of the conferences

Why authors submit

- **Knowledge dissemination**: “I want people to know about the new knowledge I discovered.” The conference promises a certain minimum level of attention one’s work gets. In other words, if you don’t publish at top conferences, nobody would read your work.
- **Feedback**: I would like people to check my results through the review process and discussion.
- **Formal goodies**: checking boxes, required to defend Ph.D., for tenure package, performance review, to put into grant report, etc.
- **Certification**. Publishing at CVPR is hard, therefore valuable.
- **Reputation-building**: Listing certain conferences on C.V. as a way of building one’s name as a scientist.

https://ducha-aiki.medium.com/what-does-it-mean-to-publish-your-scientific-paper-in-2020-9ec3a8a90fe8 https://amytabb.com

Functions of the conferences

Why authors submit

- **Knowledge dissemination**: "I want people to know about the new knowledge I discovered." The conference promises a certain minimum level of attention one's work gets. In other words, if you don't publish at top conferences, nobody would read your work.
- **Feedback**: I would like people to check my results through the review process and discussion.
- **Formal goodies**: checking boxes, required to defend Ph.D., for tenure package, performance review, to put into grant report, etc.
- **Certification**. Publishing at CVPR is hard, therefore valuable.
- **Reputation-building**: Listing certain conferences on C.V. as a way of building one's name as a scientist.
- **Networking**: meeting with peers, potential employers, etc

https://ducha-aiki.medium.com/what-does-it-mean-to-publish-your-scientific-paper-in-2020-9ec3a8a90fe8 https://amytabb.com

Functions of the conferences

Why people attend

Functions of the conferences

Why people attend

- **Prefiltering**: time is limited, so we outsource the selection of what we are reading to the reviewers.

https://ducha-aiki.medium.com/what-does-it-mean-to-publish-your-scientific-paper-in-2020-9ec3a8a90fe8 https://amytabb.com

Functions of the conferences

Why people attend

- **Prefiltering**: time is limited, so we outsource the selection of what we are reading to the reviewers.
- **Certification**: time is limited, so we outsource the quality control and result check to the reviewers. We create a basic classifier: “If the paper is published by a top-conference, it is true.”

https://ducha-aiki.medium.com/what-does-it-mean-to-publish-your-scientific-paper-in-2020-9ec3a8a90fe8 https://amytabb.com

Functions of the conferences

Why people attend

- **Prefiltering**: time is limited, so we outsource the selection of what we are reading to the reviewers.
- **Certification**: time is limited, so we outsource the quality control and result check to the reviewers. We create a basic classifier: “If the paper is published by a top-conference, it is true.”
- **Special case of certification**: for people outside the field without the basic qualifications to select work that meets basic quality guarantees.

https://ducha-aiki.medium.com/what-does-it-mean-to-publish-your-scientific-paper-in-2020-9ec3a8a90fe8 https://amytabb.com

Functions of the conferences

Why people attend

- **Prefiltering**: time is limited, so we outsource the selection of what we are reading to the reviewers.
- **Certification**: time is limited, so we outsource the quality control and result check to the reviewers. We create a basic classifier: “If the paper is published by a top-conference, it is true.”
- **Special case of certification**: for people outside the field without the basic qualifications to select work that meets basic quality guarantees.
- **Authors promise to answer our questions** (symmetrical to “attention for the author from audience,” and audience gets the guarantee that questions about the work will be answered at the talk or poster session).

https://ducha-aiki.medium.com/what-does-it-mean-to-publish-your-scientific-paper-in-2020-9ec3a8a90fe8 https://amytabb.com

Thank you for your attention

Dmytro Mishkin, PhD

ducha.aiki@gmail.com
http://dmytro.ai/
ducha_aiki

Thank you for your attention

Dmytro Mishkin, PhD

ducha.aiki@gmail.com
http://dmytro.ai/
ducha_aiki

- Computer vision researcher at CTU in Prague.

Thank you for your attention

Dmytro Mishkin, PhD

ducha.aiki@gmail.com
http://dmytro.ai/
ducha_aiki

- Computer vision researcher at CTU in Prague.

Thank you for your attention

Dmytro Mishkin, PhD

ducha.aiki@gmail.com
http://dmytro.ai/
ducha_aiki

- Computer vision researcher at CTU in Prague.
- Member of the Expert Committee on Artificial Intelligence at Ministry of Digital Transformation Of Ukraine

Thank you for your attention

Dmytro Mishkin, PhD

ducha.aiki@gmail.com
http://dmytro.ai/
ducha_aiki

- Computer vision researcher at CTU in Prague.
- Member of the Expert Committee on Artificial Intelligence at Ministry of Digital Transformation Of Ukraine

Thank you for your attention

Dmytro Mishkin, PhD

ducha.aiki@gmail.com
http://dmytro.ai/
ducha_aiki

- Computer vision researcher at CTU in Prague.
- Member of the Expert Committee on Artificial Intelligence at Ministry of Digital Transformation Of Ukraine
- Computer vision and deep learning consultant

Thank you for your attention

Dmytro Mishkin, PhD

ducha.aiki@gmail.com
http://dmytro.ai/
ducha_aiki

- Computer vision researcher at CTU in Prague.
- Member of the Expert Committee on Artificial Intelligence at Ministry of Digital Transformation Of Ukraine
- Computer vision and deep learning consultant

Thank you for your attention

Dmytro Mishkin, PhD

ducha.aiki@gmail.com
http://dmytro.ai/
ducha_aiki

- Computer vision researcher at CTU in Prague.
- Member of the Expert Committee on Artificial Intelligence at Ministry of Digital Transformation Of Ukraine
- Computer vision and deep learning consultant
- Co-author of the first Ukrainian publications at CVPR an NeurIPS

Thank you for your attention

Dmytro Mishkin, PhD

ducha.aiki@gmail.com
http://dmytro.ai/
ducha_aiki

- Computer vision researcher at CTU in Prague.
- Member of the Expert Committee on Artificial Intelligence at Ministry of Digital Transformation Of Ukraine
- Computer vision and deep learning consultant
- Co-author of the first Ukrainian publications at CVPR an NeurIPS
- CVPR 2021 Outstanding reviewer

Thank you for your attention

Dmytro Mishkin, PhD

ducha.aiki@gmail.com
http://dmytro.ai/
ducha_aiki

- Computer vision researcher at CTU in Prague.
- Member of the Expert Committee on Artificial Intelligence at Ministry of Digital Transformation Of Ukraine
- Computer vision and deep learning consultant
- Co-author of the first Ukrainian publications at CVPR an NeurIPS
- CVPR 2021 Outstanding reviewer
- Kaggle Master