Advances in Efficient Neural Network Training

Simple Models, Smarter Data

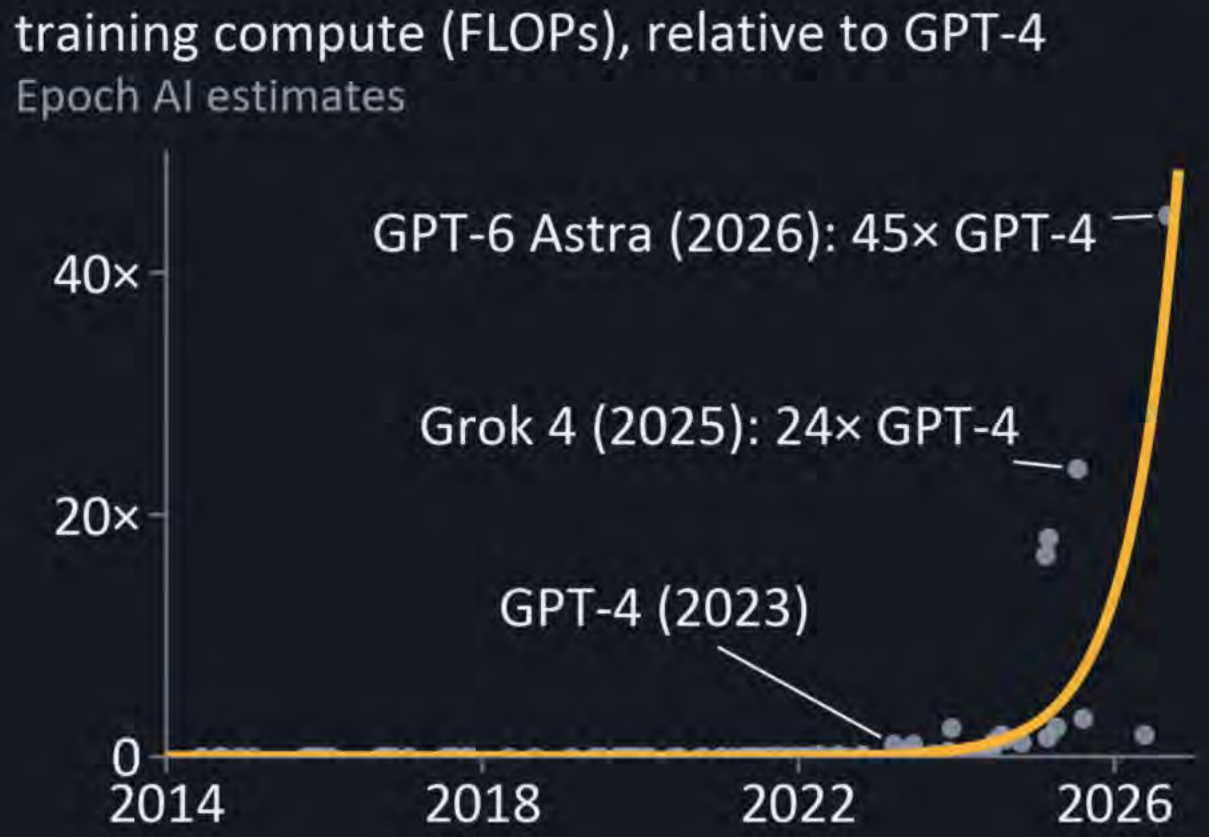
Dr. Stanislav Frolov

Smart Data & Knowledge Services (SDS), Prof. Andreas Dengel
German Research Center for Artificial Intelligence (DFKI)

stanislav.frolov@dfki.de

Trilateral AI Symposium · Tokyo · October 2026

Training compute grows up to 5× a year



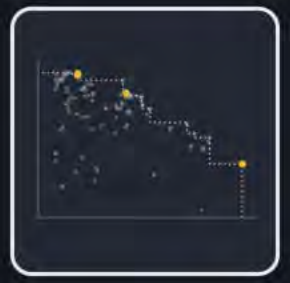
Bigger models need more data
=> bigger bills
=> fewer labs at the frontier

Good data is slow and expensive
=> need for higher data efficiency

Jaime Sevilla, Edu Roldán
Training Compute of Frontier AI Models Grows by 4–5x per Year, **Epoch AI 2024**

Epoch AI, *Data on Notable AI Models*, **2026**

Talk Outline



1 Which transformer is really the most efficient?



2 Can we get more out of our images and labels?



3 Can generated images replace real images?

1 Which transformer is really the most efficient?



=> use ViT variants on ImageNet as standard benchmark

Tobias Nauen, Sebastian Palacio, Federico Raue, Andreas Dengel
Which Transformer to Favor: A Comparative Analysis of Efficiency in Vision Transformers, **WACV 2025**

Benchmark 45+ ‘efficient’ models

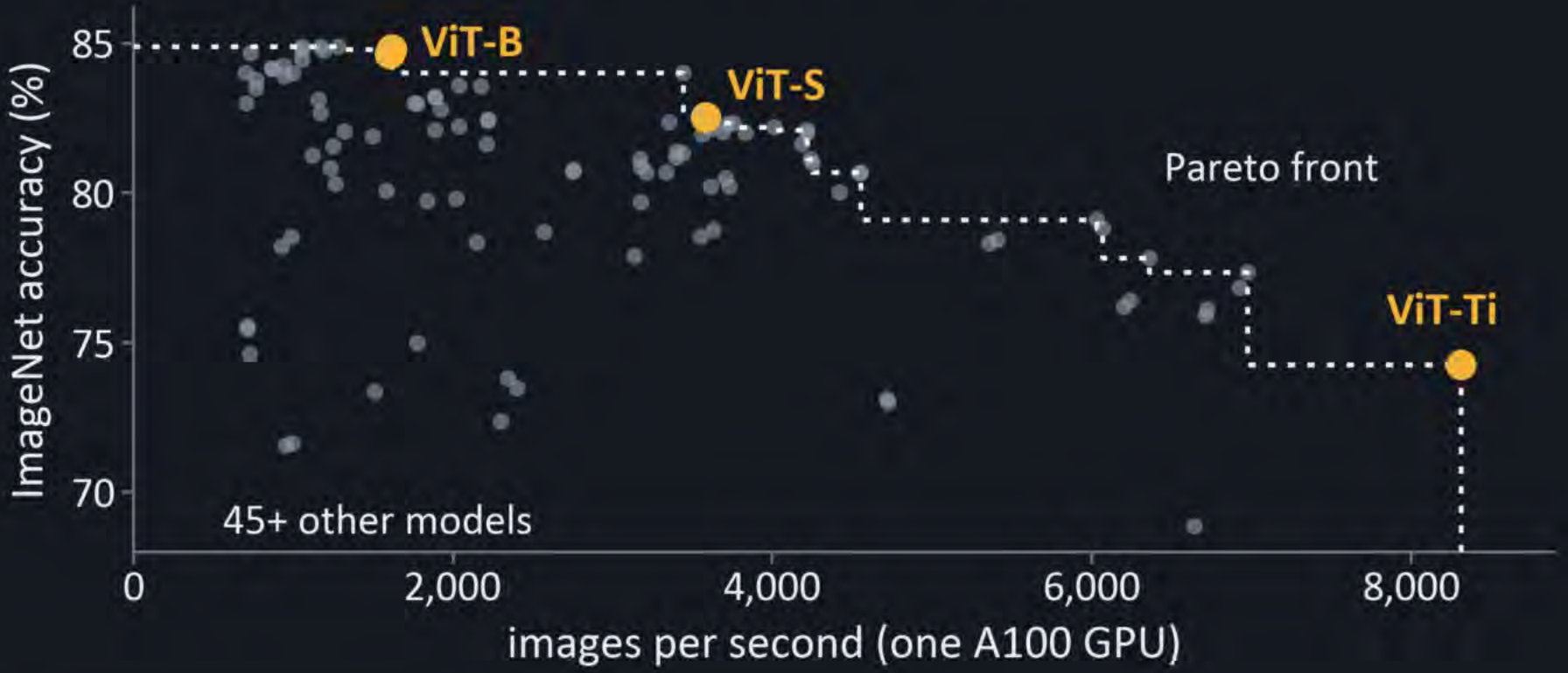


same data
same GPUs
same training recipe
=> fair speed and memory comparison

← **the plain ViT**

Tobias Nauen, Sebastian Palacio, Federico Raue, Andreas Dengel
Which Transformer to Favor: A Comparative Analysis of Efficiency in Vision Transformers, **WACV 2025**

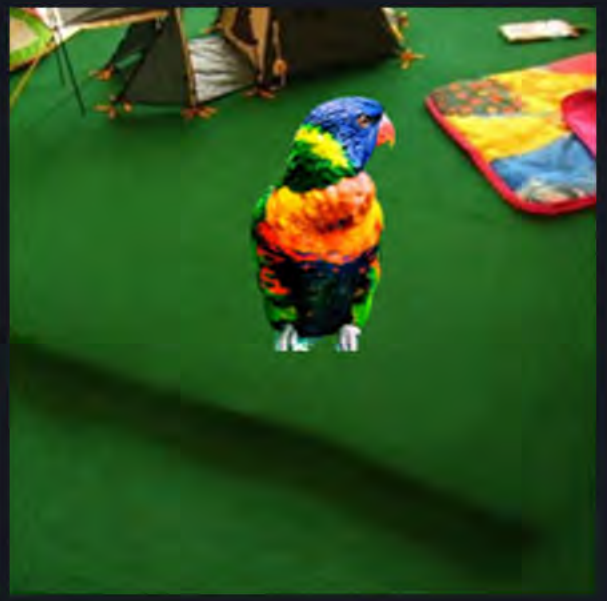
The plain ViT is still Pareto-optimal



1 Use simple models that scale

Tobias Nauen, Sebastian Palacio, Federico Raue, Andreas Dengel
Which Transformer to Favor: A Comparative Analysis of Efficiency in Vision Transformers, **WACV 2025**

2 Can we get more out of our images and labels?



parrot?



goose or cat?

Tobias Nauen, Brian Moser, Federico Raue, Stanislav Frolov, Andreas Dengel
ForAug: Mitigating Biases in Image Classification via Controlled Image Compositions, **Under Review**

Tobias Nauen, Stanislav Frolov, Federico Raue, Brian Moser, Andreas Dengel
OA-CutMix: Correcting the Label Bias of CutMix, **ICANN 2026**

A parrot, or a sleeping bag?



original image



sleeping bag inpainted, parrot pasted in

ViT-B, trained on ImageNet

sleeping bag
69%

Tobias Nauen, Brian Moser, Federico Raue, Stanislav Frolov, Andreas Dengel
ForAug: Mitigating Biases in Image Classification via Controlled Image Compositions, **Under Review**

ImageNet's fish images come with a fisherman

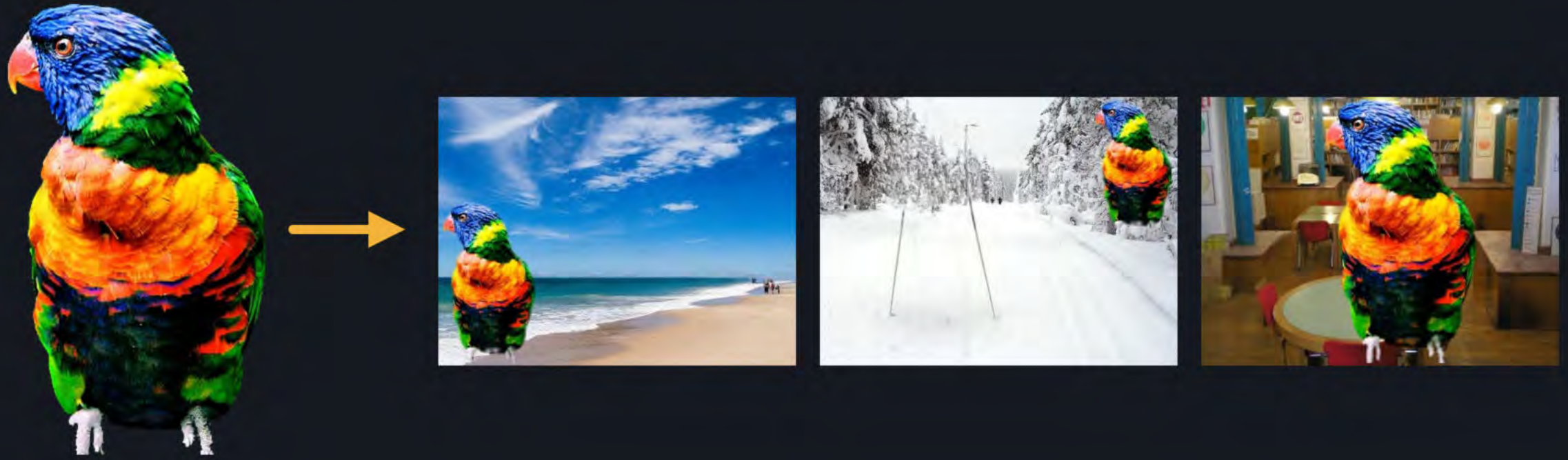


Dataset biases such as

- **objects in the center**
- **at a typical size**
- **in their usual surroundings**

Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, Li Fei-Fei
ImageNet: A Large-Scale Hierarchical Image Database, **CVPR 2009**

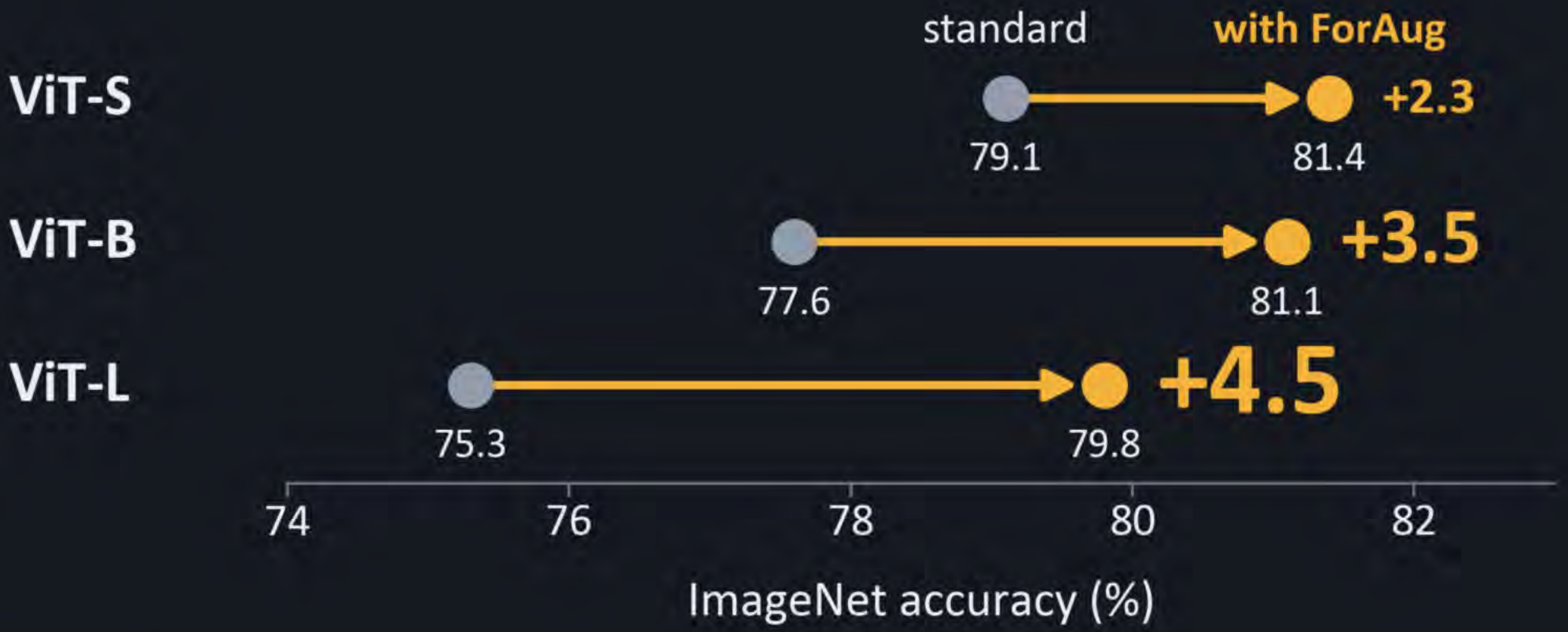
Segment, cut out, clean up, recombine



+ clean backgrounds
+ control of place, scale and scene

Tobias Nauen, Brian Moser, Federico Raue, Stanislav Frolov, Andreas Dengel
ForAug: Mitigating Biases in Image Classification via Controlled Image Compositions, **Under Review**

Improvements scale with the model size



Tobias Nauen, Brian Moser, Federico Raue, Stanislav Frolov, Andreas Dengel
ForAug: Mitigating Biases in Image Classification via Controlled Image Compositions, **Under Review**

The same parrot, after ForAug



ViT-B, trained on ImageNet

~~**sleeping bag 69%**~~

ViT-B, trained with ForAug

parrot 81%

Tobias Nauen, Brian Moser, Federico Raue, Stanislav Frolov, Andreas Dengel

ForAug: Mitigating Biases in Image Classification via Controlled Image Compositions, **Under Review**

Data augmentation: more from the same images

dfki ai

classic



original



flip



crop



color



rotate



blur



grayscale

mixing two images



MixUp



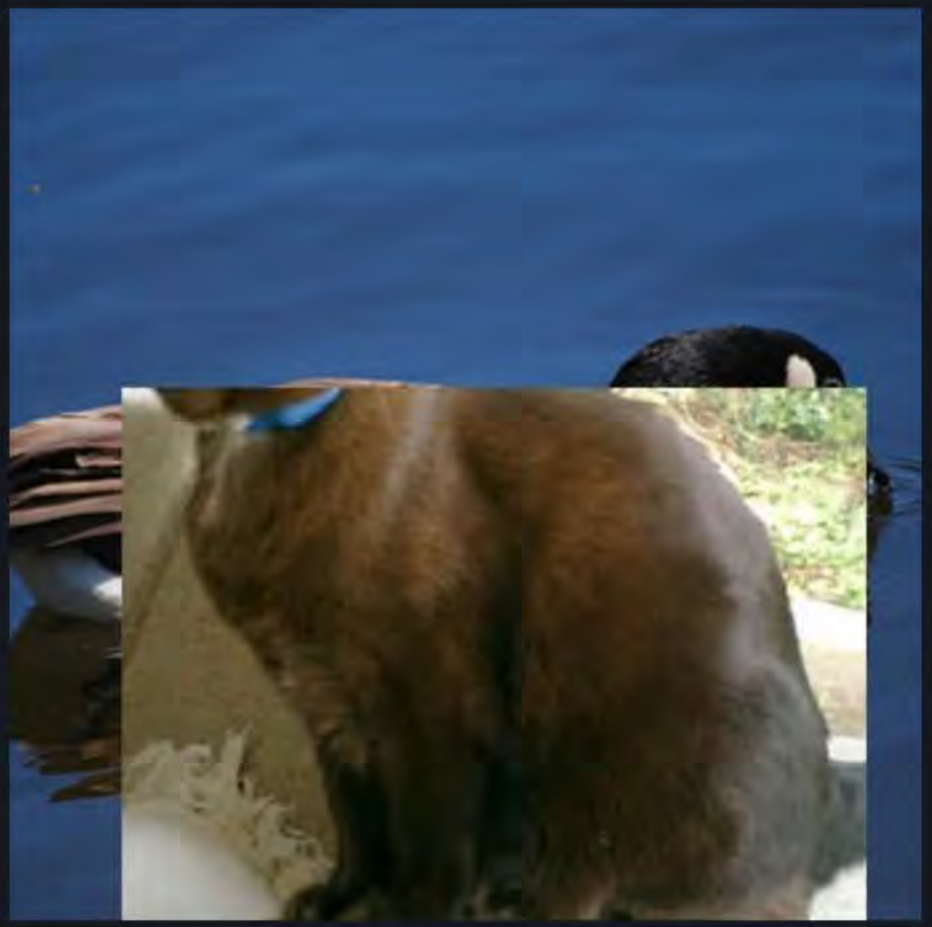
CutMix

CutMix: simple, works, used almost everywhere

Hongyi Zhang, Moustapha Cisse, Yann N. Dauphin, David Lopez-Paz
Mixup: Beyond Empirical Risk Minimization, **ICLR 2018**

Sangdoo Yun, Dongyoon Han, Seong Joon Oh, Sanghyuk Chun, Junsuk Choe, Youngjoon Yoo
CutMix: Regularization Strategy to Train Strong Classifiers with Localizable Features, **ICCV 2019**

CutMix: paste a patch, mix the labels by size



CutMix label

48% cat

52% goose

visible: mostly cat

Sangdoo Yun, Dongyoon Han, Seong Joon Oh, Sanghyuk Chun, Junsuk Choe, Youngjoon Yoo
CutMix: Regularization Strategy to Train Strong Classifiers with Localizable Features, **ICCV 2019**

CutMix images produce ghost labels



17%

of CutMix augmented images name an object that isn’t really visible

Tobias Nauen, Stanislav Frolov, Federico Raue, Brian Moser, Andreas Dengel
OA-CutMix: Correcting the Label Bias of CutMix, **ICANN 2026**

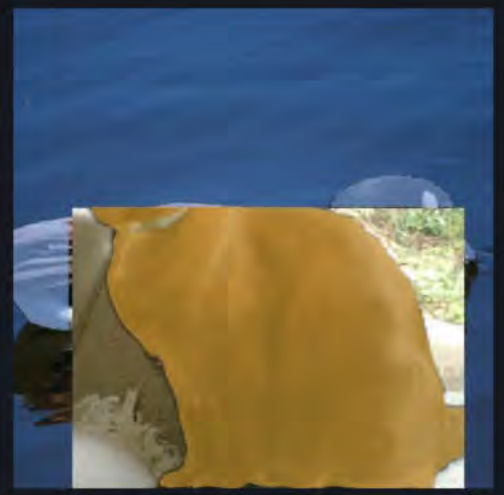
Object-Aware (OA)-CutMix




1 find each object (done offline)



2 mix as usual



3 label what is visible

82% cat · **18% goose**

+ fixed label
+ same training speed

Tobias Nauen, Stanislav Frolov, Federico Raue, Brian Moser, Andreas Dengel
OA-CutMix: Correcting the Label Bias of CutMix, **ICANN 2026**

Best accuracy, at CutMix speed

dfki ai

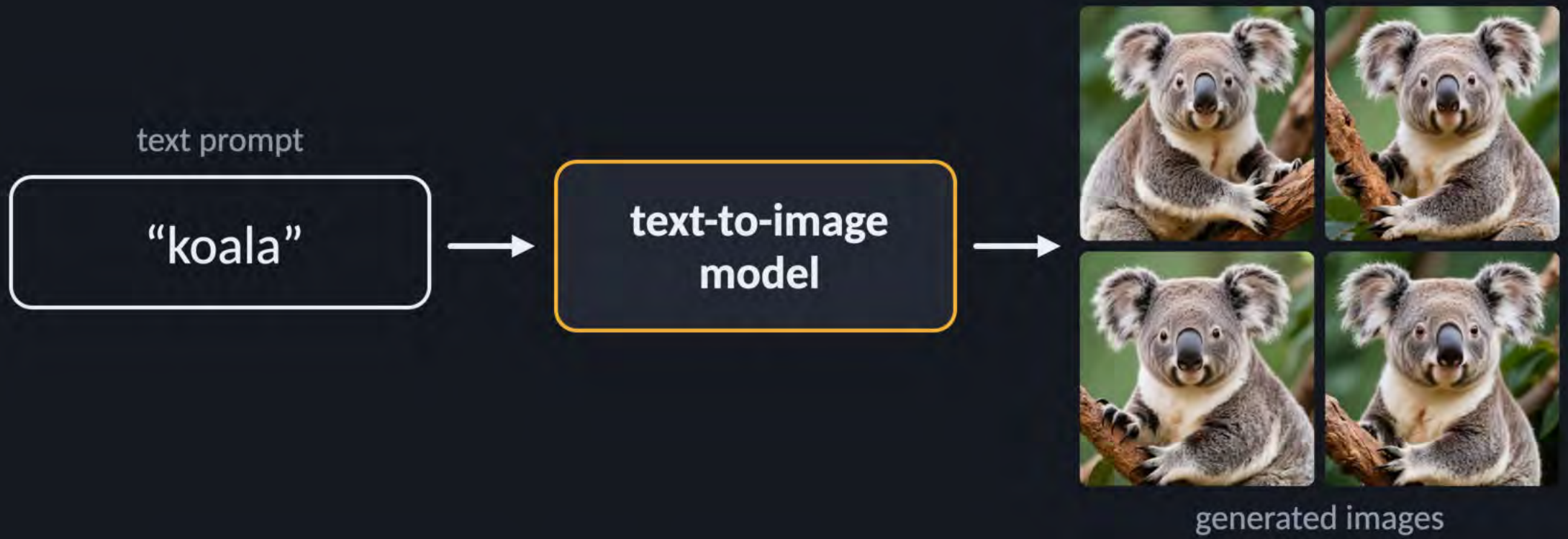
10+ mixing methods over 3 datasets and 2 models

	ViT (DeiT)	ResNet	mixing time
CutMix	74.0	75.3	→ **0.6 ms**
Puzzle Mix	72.9	76.2	——————→ **301 ms**
OA-CutMix	**75.9**	**76.3**	→ **1.0 ms**

2 **Look deeply into your data and labels**

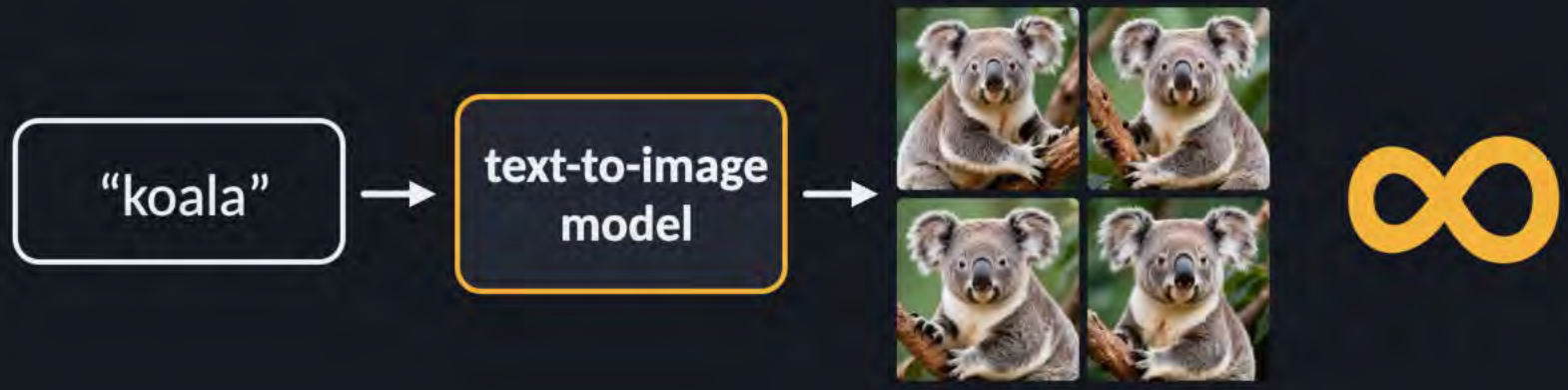
Tobias Nauen, Stanislav Frolov, Federico Raue, Brian Moser, Andreas Dengel
OA-CutMix: Correcting the Label Bias of CutMix, **ICANN 2026**

3 Can generated images replace real images?



Krzysztof Adamkiewicz, Brian Moser, Stanislav Frolov, Tobias Nauen, Federico Raue, Andreas Dengel
When Pretty Isn’t Useful: Investigating Why Modern Text-to-Image Models Fail as Reliable Training Data Generators, **CVPR 2026**

The promise: unlimited training data



- on demand
- already labeled
- no privacy issues

Image generators got much better, 2022 → 2025

dfki ai



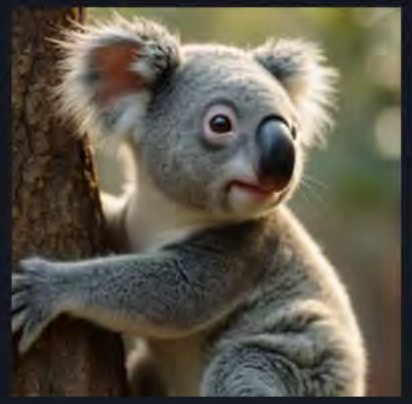
SD 1.5
2022 · GenEval 0.43



SD 2.1
2022 · GenEval 0.50



SD 3 Medium
2024 · GenEval 0.62



FLUX.1-dev
2024 · GenEval 0.66



Qwen-Image
2025 · GenEval 0.87

sharper, more realistic, better at following the prompt


Labs build better image generators
So we get better training data, right?
Better training data, right?

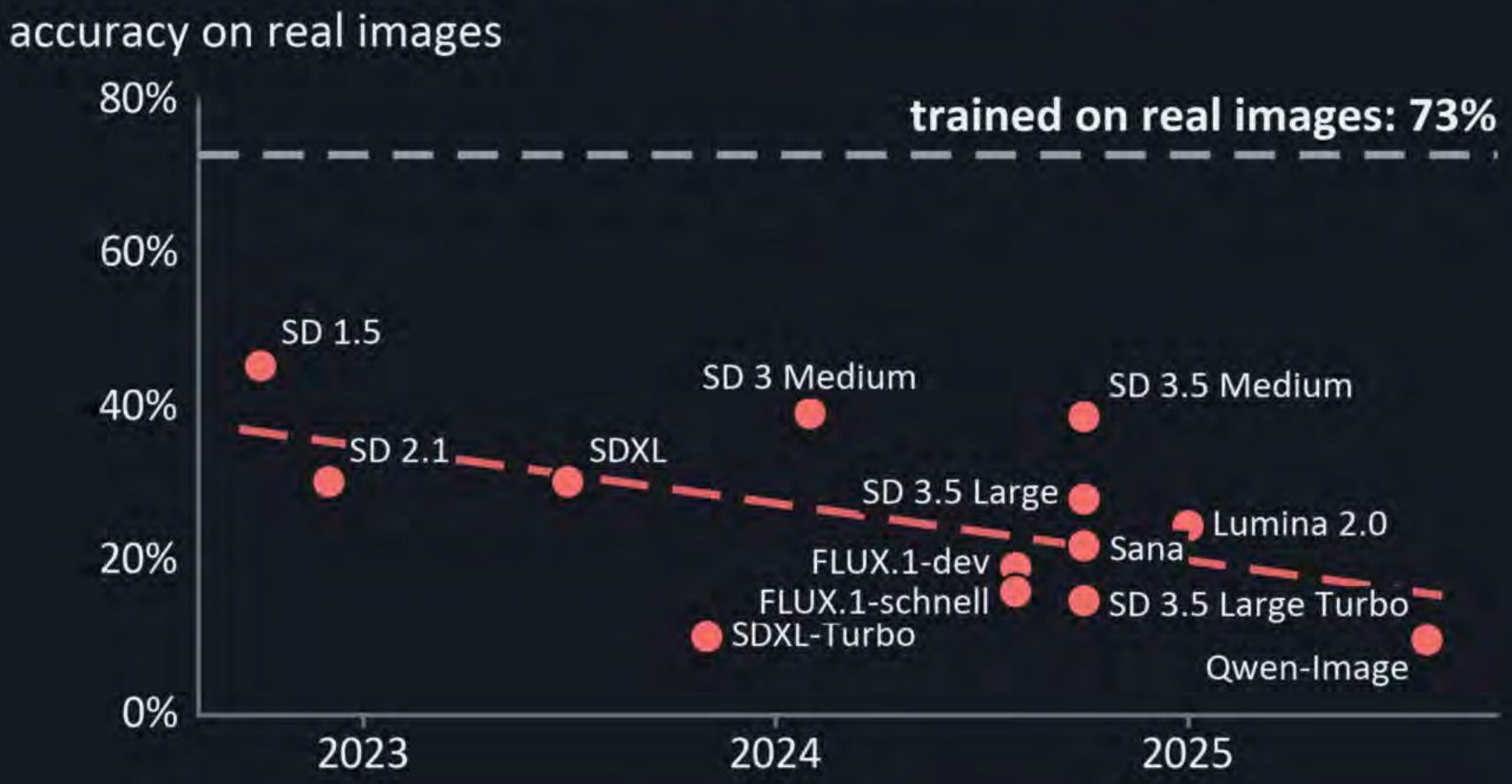
Setup: train on generated, test on real images



Benchmark +10 image generation models

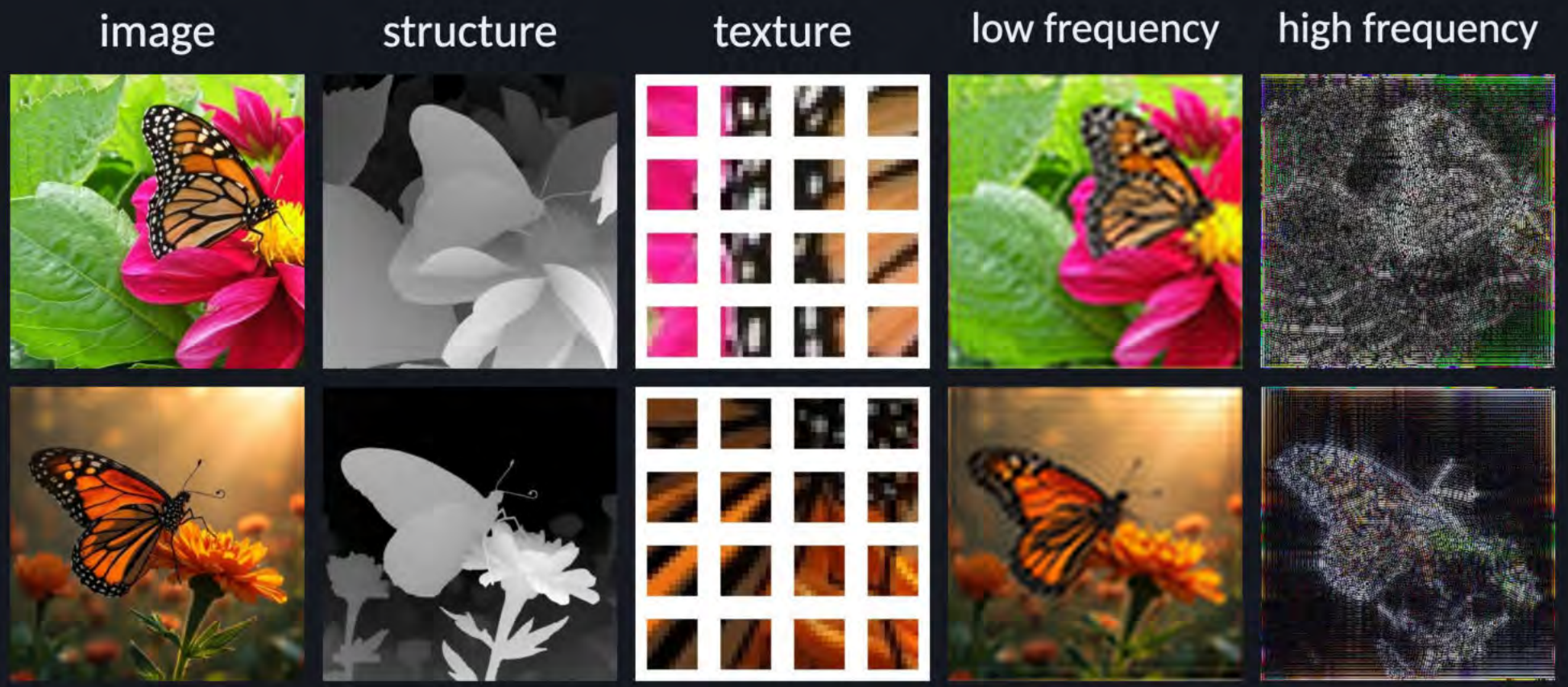
Krzysztof Adamkiewicz, Brian Moser, Stanislav Frolov, Tobias Nauen, Federico Raue, Andreas Dengel
When Pretty Isn’t Useful: Investigating Why Modern Text-to-Image Models Fail as Reliable Training Data Generators, **CVPR 2026**

Newer generators, worse training data



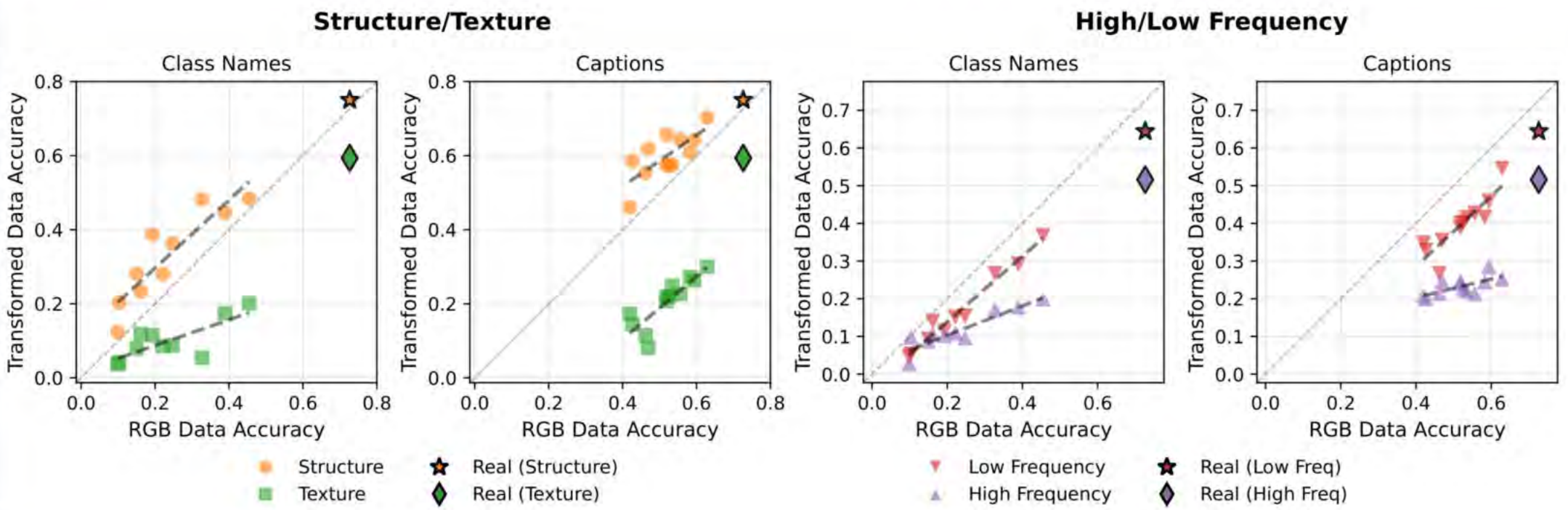
Krzysztof Adamkiewicz, Brian Moser, Stanislav Frolov, Tobias Nauen, Federico Raue, Andreas Dengel
When Pretty Isn't Useful: Investigating Why Modern Text-to-Image Models Fail as Reliable Training Data Generators, **CVPR 2026**

Why? Decompose for further analysis



Krzysztof Adamkiewicz, Brian Moser, Stanislav Frolov, Tobias Nauen, Federico Raue, Andreas Dengel
When Pretty Isn't Useful: Investigating Why Modern Text-to-Image Models Fail as Reliable Training Data Generators, **CVPR 2026**

What is missing: texture and high frequencies



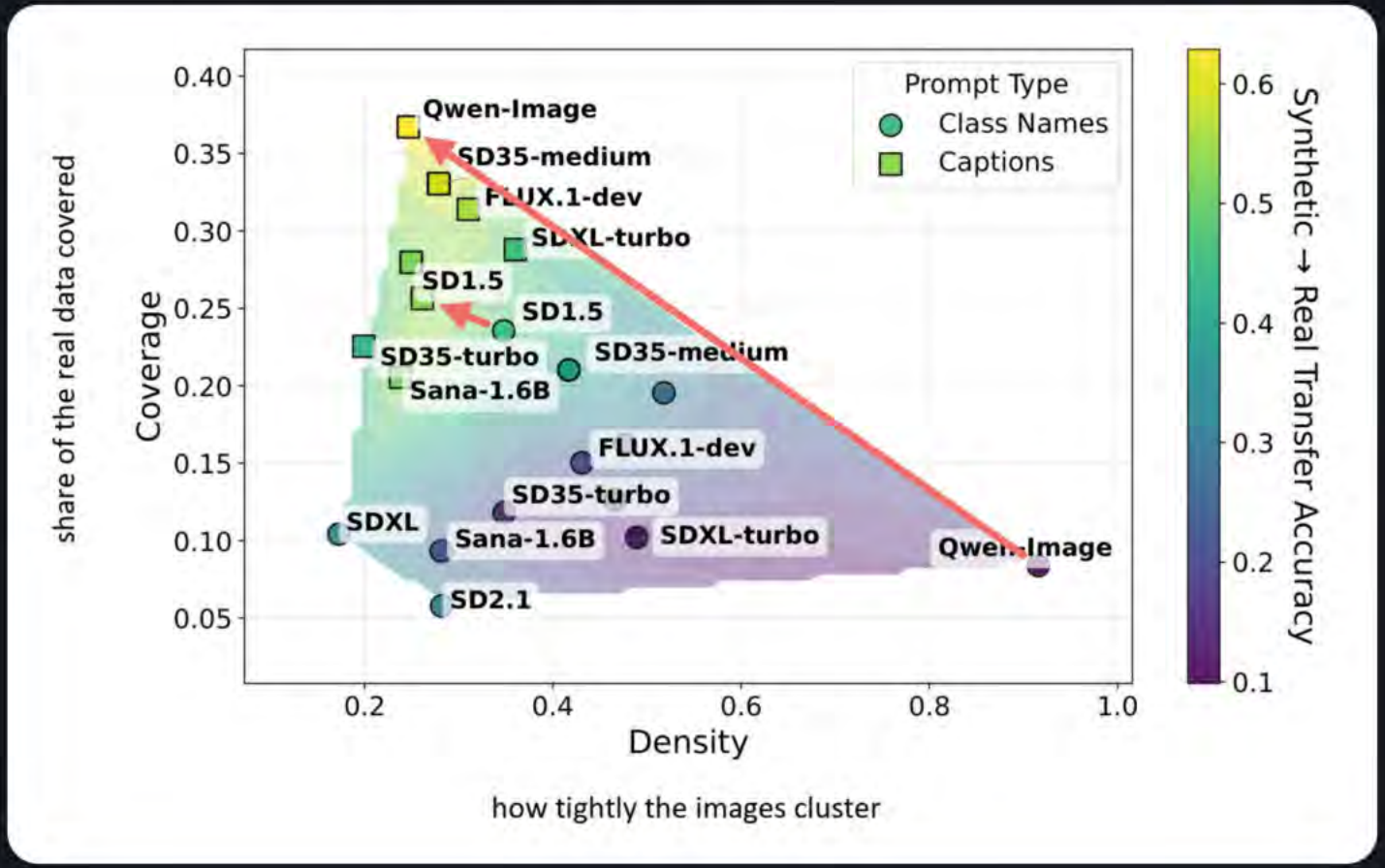
+ structure and low frequencies

- texture and high frequencies

Krzysztof Adamkiewicz, Brian Moser, Stanislav Frolov, Tobias Nauen, Federico Raue, Andreas Dengel

When Pretty Isn't Useful: Investigating Why Modern Text-to-Image Models Fail as Reliable Training Data Generators, **CVPR 2026**

Prettier, but less diverse



- Newest generators produce dense but narrow images
- Detailed prompts add coverage
- Newer models rely on good prompting

Krzysztof Adamkiewicz, Brian Moser, Stanislav Frolov, Tobias Nauen, Federico Raue, Andreas Dengel
When Pretty Isn't Useful: Investigating Why Modern Text-to-Image Models Fail as Reliable Training Data Generators, **CVPR 2026**

Generated images alone are far from real data

73%
trained on real images

46%
best of 13 generators

63%
real image prompts

Build generators for usefulness, not only for aesthetics

3 **Don’t rely on synthetic data alone**

Krzysztof Adamkiewicz, Brian Moser, Stanislav Frolov, Tobias Nauen, Federico Raue, Andreas Dengel
When Pretty Isn’t Useful: Investigating Why Modern Text-to-Image Models Fail as Reliable Training Data Generators, **CVPR 2026**

Takeaways

1. Use simple models that scale
2. Look deeply into your data and labels
3. Don’t rely on synthetic data alone

Thank you!

Dr. Stanislav Frolov · stanislav.frolov@dfki.de


Tobias Nauen


Krzysztof Adamkiewicz


Brian Moser


Sebastian Palacio


Federico Raue


Andreas Dengel

More Information