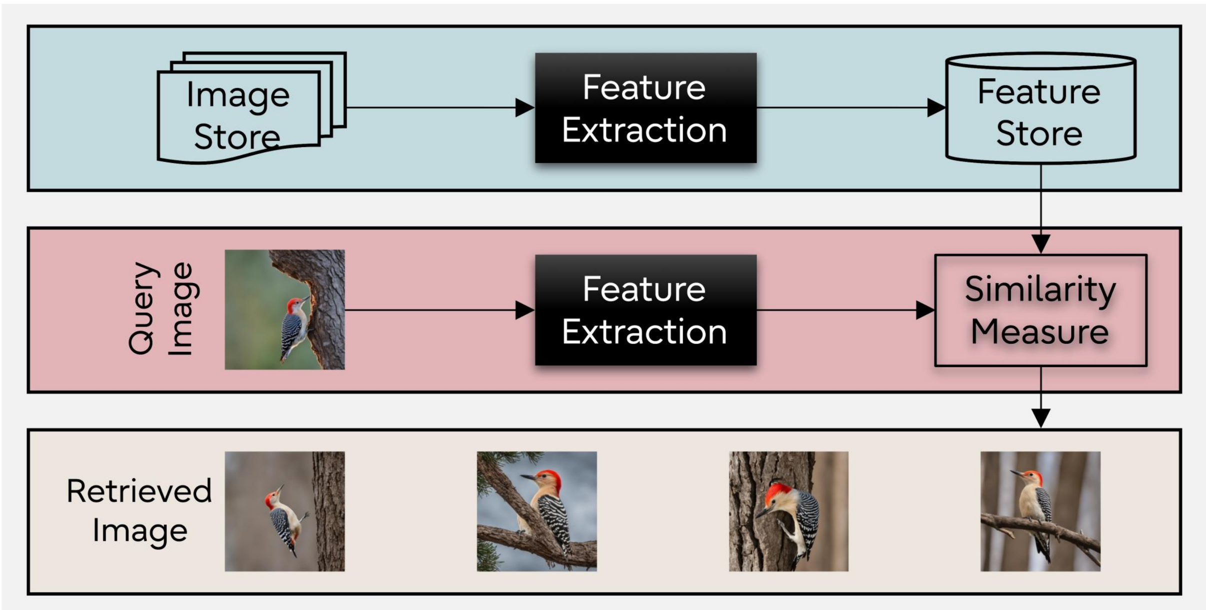
ATLANTIS: A Framework for Automated Targeted Language-guided Augmentation Training for Robust Image Search

Inderjeet Singh(inderjeet.singh@fujitsu.com), Roman Vainshtein, Alon Zolfi, Asaf Shabtai, Tu Bui, Jonathan Brokman, Omer Hofman, Fumiyoshi Kasahara, Kentaro Tsuji, Hisashi Kojima*

INTRODUCTION

Content-Based Image Retrieval

- Content-Based Image Retrieval (CBIR) is a technique for finding images in a database based on their visual content by using deep metric learning (DML).



Challenges in CBIR Systems

- **Generalization issues** [Neyshabur+, 2017]: due to domain gaps and class imbalances.
- **Overfitting** [Neyshabur+, 2017]
- **Susceptibility to adversarial attacks** [Xie+, 2020].

Available training data

Traditional cropping, rotation, flipping, and colour transformation-based **augmentation**: less control over content distribution.

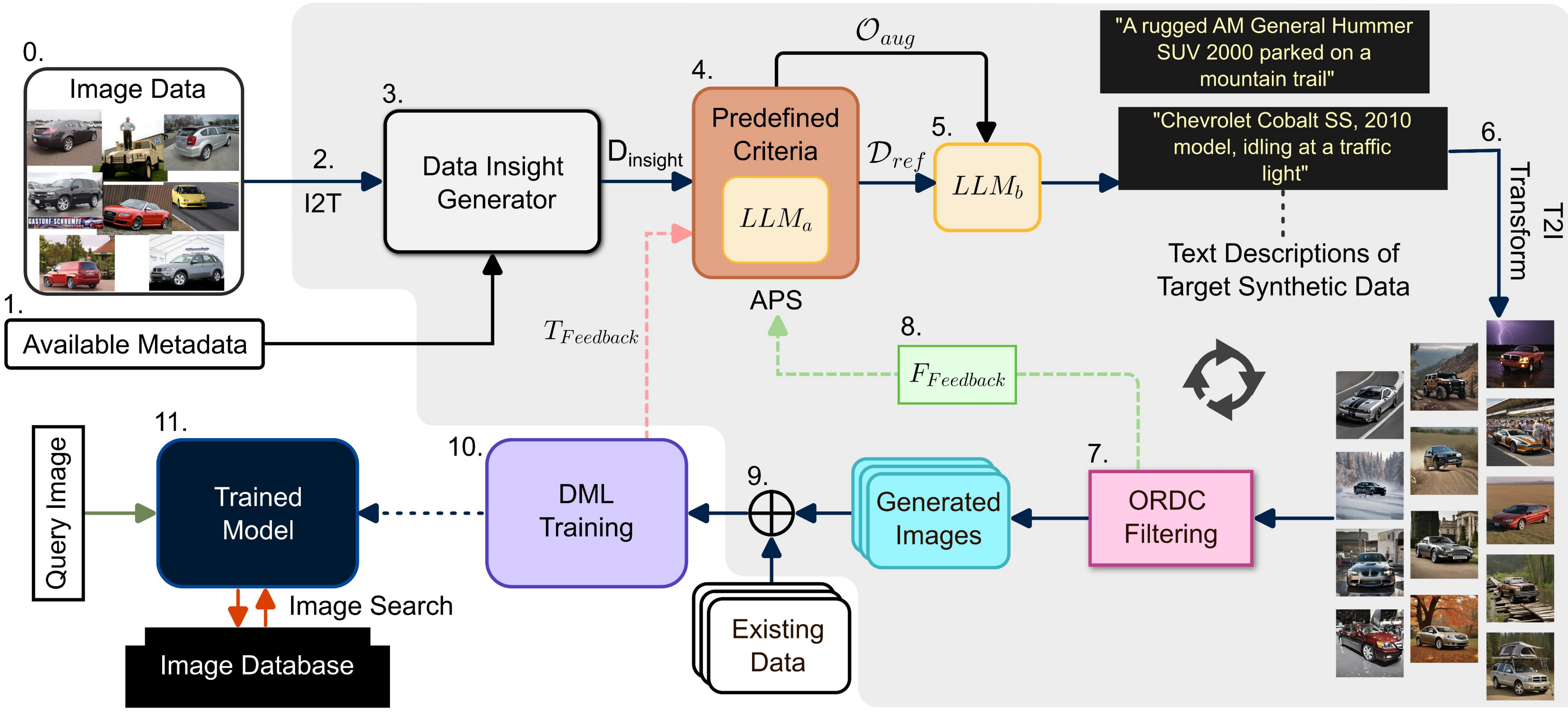
Generative augmentation: improved content distribution.

Existing Data Augmentation Works

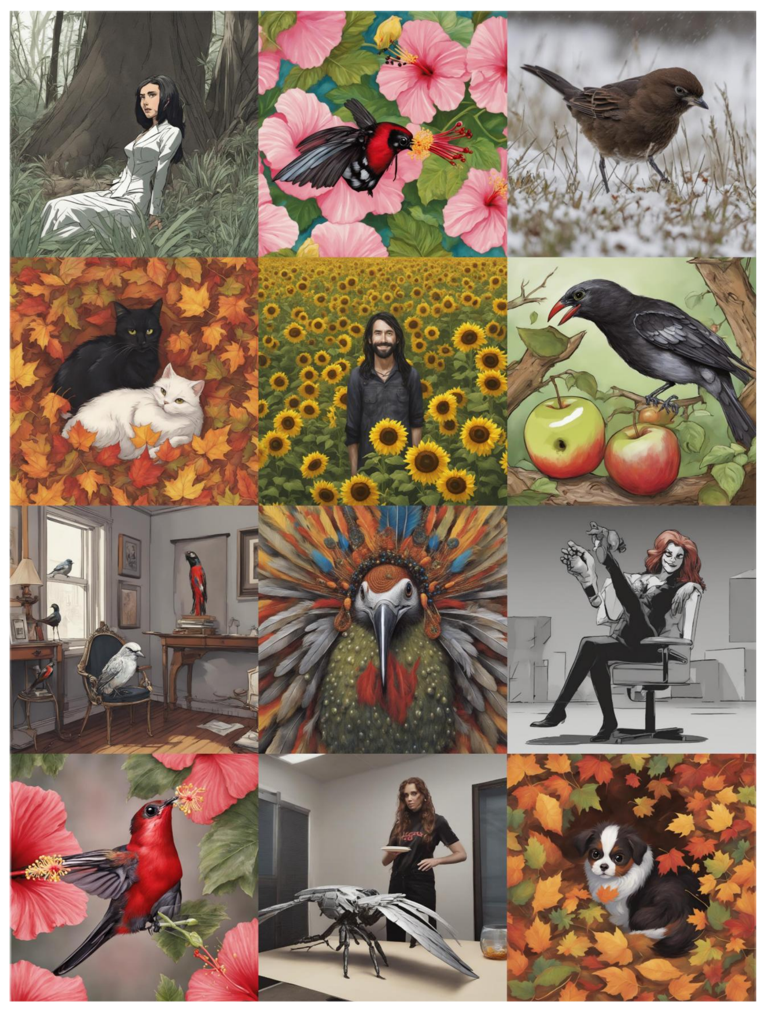
- **Traditional Image Augmentation** [Perez+, 2017; Shorten+, 2019]: **cropping, rotation, flipping,** etc. However, they are **limited in their ability to change the content distribution.**
- **Generative Adversarial Networks (GANs)** [Anotoniou+, 2017; Bowles+ 2018]. **Computationally expensive, difficult to scale,** and **less control over generations.**

ATLANTIS FRAMEWORK

ATLANTIS employs a **multimodal approach** for **targeted synthetic data augmentation** that directly **identifies weaknesses and gaps in the available training data and addresses them efficiently**.



An overview of the ATLANTIS



Semantic noise from SDXL for *'Chuck-will's-Widow'* bird species.

KEY RESULTS

Data	PD	DINO$_H$ R@1	DINO$_H$ R@2	DINO$_H$ R@4	AD	Δ^*	**A-DINO$_H$** R@1	**A-DINO$_H$** R@2	**A-DINO$_H$** R@4
$\mathcal{X}_{cub}$	A_f	64.2	75.6	84.4	A_s,A_a	1	**68.4**	**78.4**	**86.3**
	A_s	69.2	78.1	85.5	A_a,A_f	1	**72.8**	**81.6**	**88.3**
	A_a	63.1	74.3	83.0	A_f,A_s	1	**66.9**	**77.6**	**85.8**
$\mathcal{X}_{cars}$	B_s	70.3	79.2	86.8	B_c,B_p	1.5	**75.4**	**84.7**	**91.0**
	B_c	66.9	76.1	83.9	B_p,B_s	1.5	**75.8**	**84.2**	**90.4**
	B_p	56.9	68.6	78.2	B_s,B_c	1.5	**75.2**	**84.5**	**90.9**
		ViT$_H$					**A-ViT$_H$**		
$\mathcal{X}_{cub}$	A_f	78.3	87.0	92.4	A_s,A_a	1	**79.7**	**88.0**	**92.9**
	A_s	79.1	87.5	92.4	A_a,A_f	1	**81.7**	**88.4**	**93.1**
	A_a	77.7	86.5	92.2	A_f,A_s	1	**79.9**	**87.5**	**92.8**
$\mathcal{X}_{cars}$	B_s	65.1	76.2	84.6	B_c,B_p	1.5	**72.8**	**82.5**	**89.7**
	B_c	62.7	73.3	81.7	B_p,B_s	1.5	**74.1**	**82.8**	**90.0**
	B_p	51.0	63.1	73.8	B_s,B_c	1.5	**71.1**	**81.6**	**89.2**

- Improvements under all **domain-scarce** and **class-imbalanced (particularly zero-shot)** scenarios.

CONCLUSION & FUTURE WORK

ATLANTIS, through the targeted synthetic data augmentation in CBIR, achieved:

- **Improved generalization** in **data-scarce scenarios**.
- **Enhanced adversarial robustness, particularly against attacks with imperceptible noise levels.**
- **Competitive performance on standard CBIR benchmarks.**

Future Work:

- Improving computational efficiency and stability.
- Enhanced ethical considerations: Future development could integrate a blacklist of objects and domains into ORDC for more controllable synthesis.

Code and related data is available at: