Detection of AI-generated images

Lucas SALAND
Supervisor: Patrick Bas

Université de Lille

CRIStAL
Centre de Recherche en Informatique, Signal et Automatique de Lille

Lille, France

August 6, 2024

1. AI image generation
   - GAN
   - Diffusion model
2. AI-generated images detection
   - CLIP
   - Impact of JPEG compression
   - Generators diversity and neural network
   - Pair training and fine-tuning
   - DoubleCLIP: Another approach to work with pairs
   - Tip-Adapter
3. Conclusion

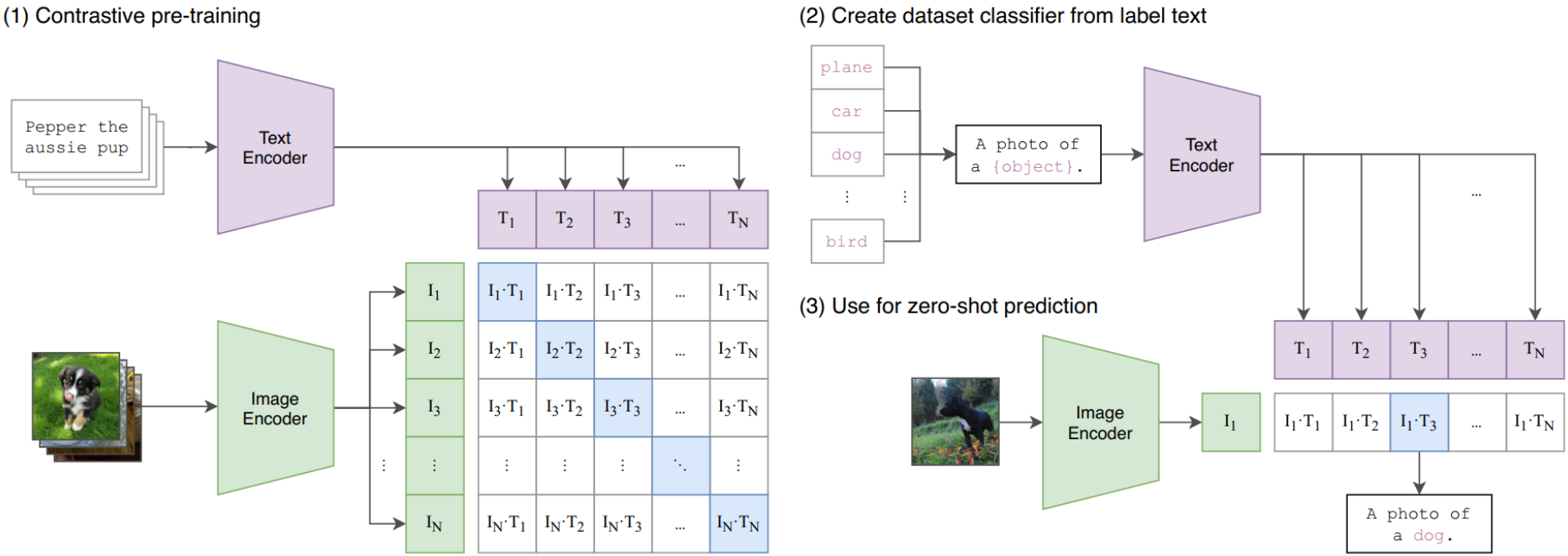
1. AI image generation
   - GAN
   - Diffusion model
2. AI-generated images detection
3. Conclusion

1 AI image generation

2 AI-generated images detection
- CLIP
- Impact of JPEG compression
- Generators diversity and neural network
- Pair training and fine-tuning
- DoubleCLIP: Another approach to work with pairs
- Tip-Adapter

3 Conclusion

CLIP


(1) Contrastive pre-training
Pepper the aussie pup
Text Encoder
Image Encoder
(2) Create dataset classifier from label text
plane
car
dog
bird
A photo of a {object}.
Text Encoder
(3) Use for zero-shot prediction
Image Encoder
A photo of a dog.

CLIP for detection

Method proposed in:[1]

[1] Cozzolino et al., *Raising the Bar of AI-generated Image Detection with CLIP.*

CLIP for detection

Method proposed in:[1]

- Build a dataset of pairs of real and generated images

[1] Cozzolino et al., *Raising the Bar of AI-generated Image Detection with CLIP*.

CLIP for detection

Method proposed in:[1]

- Build a dataset of pairs of real and generated images
- Extract the CLIP features

[1] Cozzolino et al., *Raising the Bar of AI-generated Image Detection with CLIP*.

CLIP for detection

Method proposed in:[1]

- Build a dataset of pairs of real and generated images
- Extract the CLIP features
- Train a SVM on these features

[1] Cozzolino et al., *Raising the Bar of AI-generated Image Detection with CLIP*.

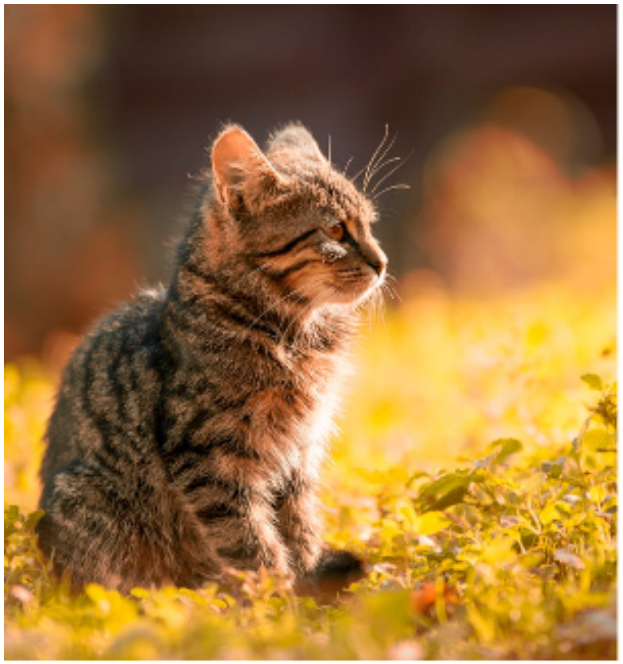
ELSA dataset

An image of 025 Greece Hellenic Air Force Lockheed Martin F16D Fighting Falcon, Medium shot, Ring light, At night, Wideangle lens

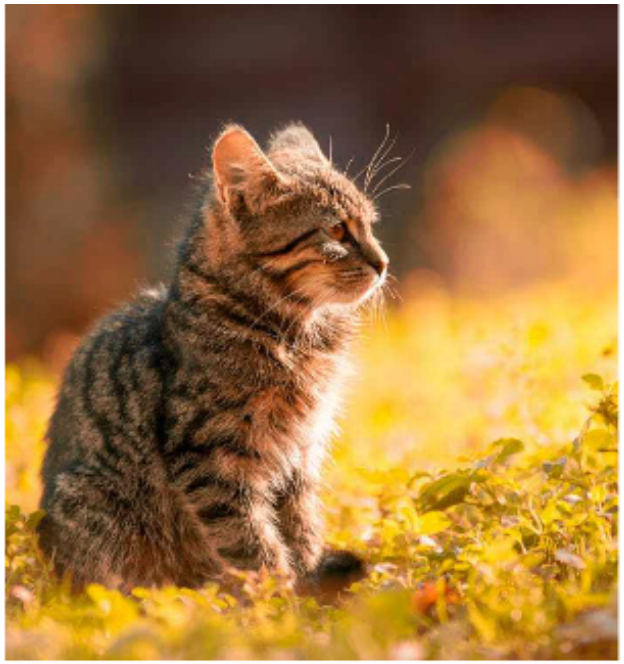
Real

Generated

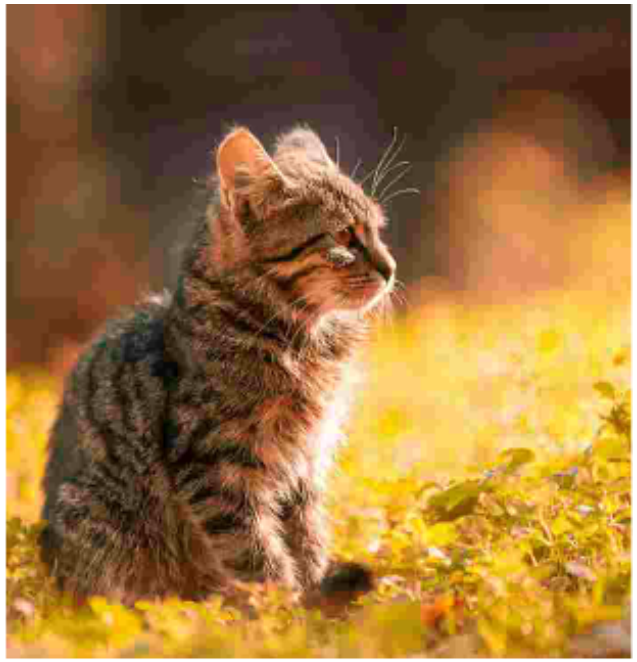
JPEG compression's impact



Original image



JPEG quality 40 (low)



JPEG quality 10 (very low)

JPEG compression's impact

	Test			
	quality	40	65	90
Train	40	0.9813	0.9730	0.9797
	65	0.9450	0.9825	0.9830
	90	0.7744	0.8581	0.9925

Table 1: Accuracy of binary classification for pairs of qualtity factors for train and test.

Generators diversity and neural network

Synthbuster:

- 9 generators
- 1000 images per generator

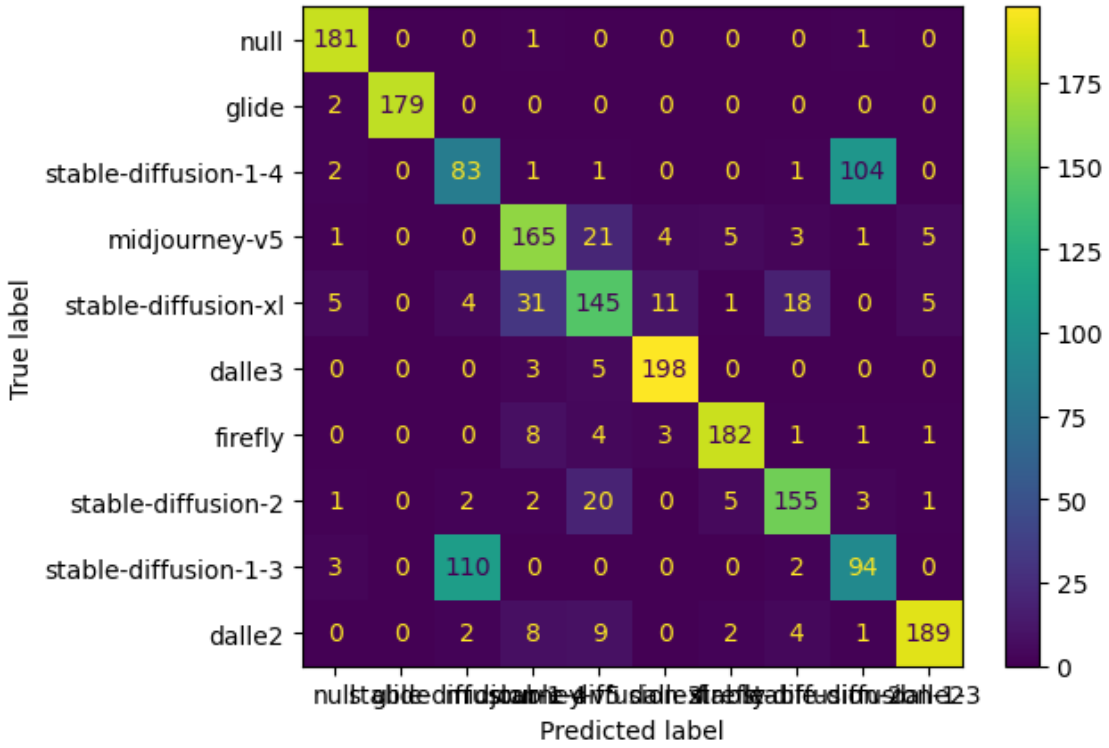
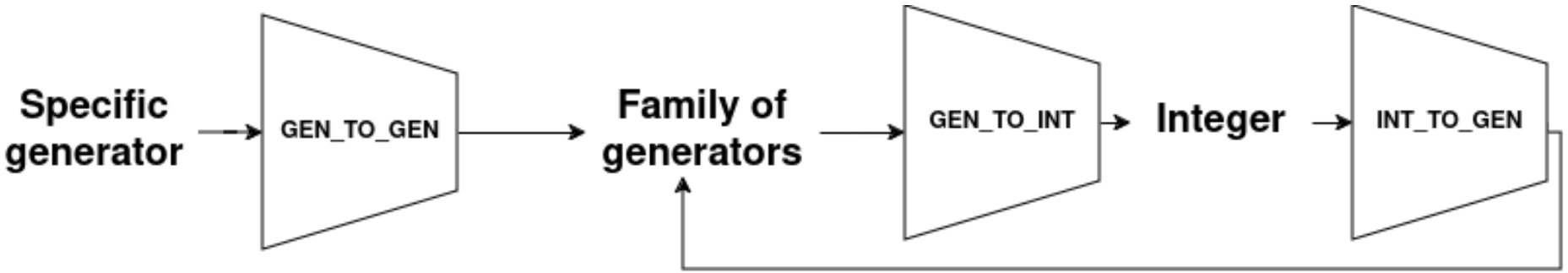


Figure 1: Confusion matrix for a SVM multiclass-classifier trained and tested on synthbuster.

Generators diversity and neural network

Multi-class classifier for binary classification:



1 AI image generation

2 AI-generated images detection

3 Conclusion