On the Effects of Artificial Data Modification

Antonia Marcu

Adam Prügel-Bennett

Why Data Modification?

Why Data Modification?

- Enhancing data sets

Why Data Modification?

- Enhancing data sets
- Evaluating models

Why Data Modification?

- Enhancing data sets
- Evaluating models
- Understanding learnt representations

What are the Implicit Assumptions?

What are the Implicit Assumptions?

Side-effects

What are the Implicit Assumptions?

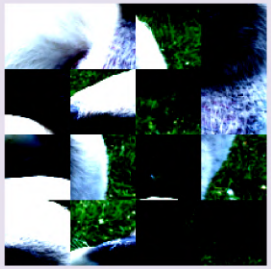
Side-effects

Evaluation

What are the Implicit Assumptions?

Side-effects

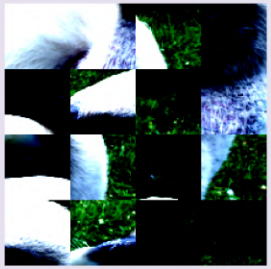
Evaluation

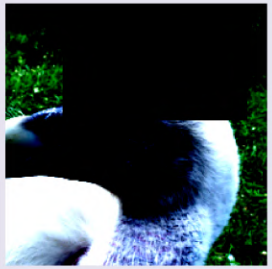


What are the Implicit Assumptions?

Side-effects

Evaluation



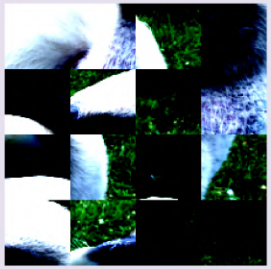


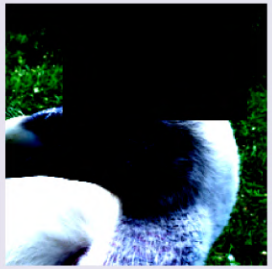
What are the Implicit Assumptions?

Side-effects

Evaluation

Training



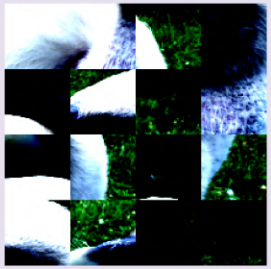


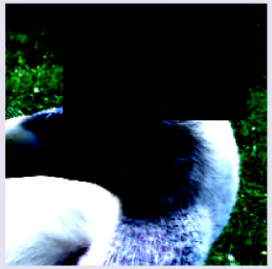
What are the Implicit Assumptions?

Side-effects

Evaluation

Training



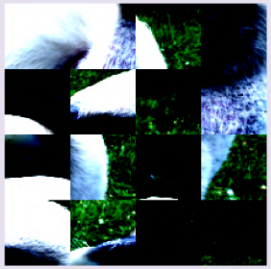


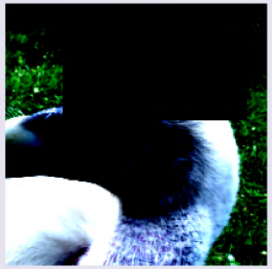
What are the Implicit Assumptions?

Side-effects

Evaluation

Training



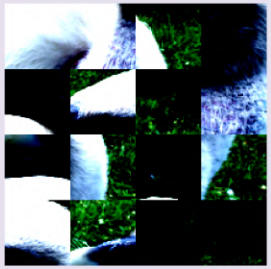


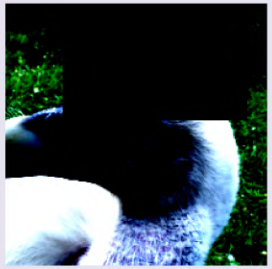
What are the Implicit Assumptions?

Side-effects

Evaluation

Training



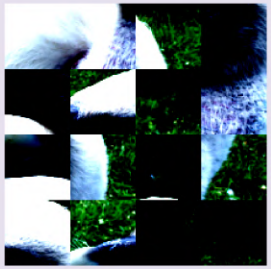


What are the Implicit Assumptions?

Side-effects

Evaluation

Training

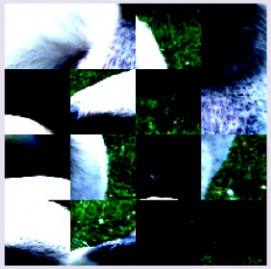


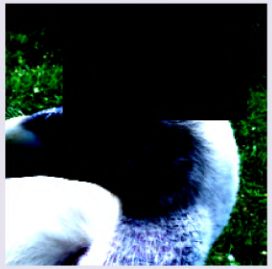
What are the Implicit Assumptions?

Side-effects

Evaluation

Training



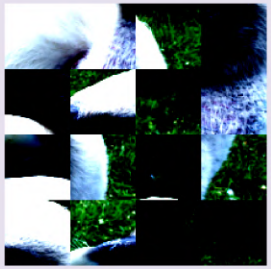


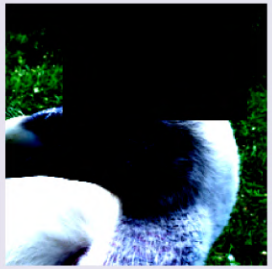
What are the Implicit Assumptions?

Side-effects

Evaluation

Training





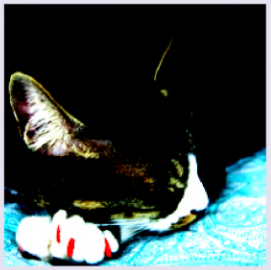
Side-effects: Evaluating

Side-effects: Evaluating

- Are artefacts associated with a class?

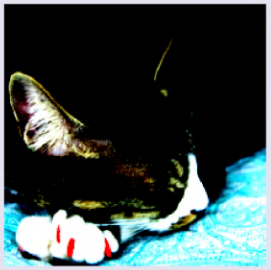
Side-effects: Evaluating

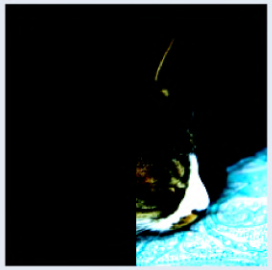
- Are artefacts associated with a class?



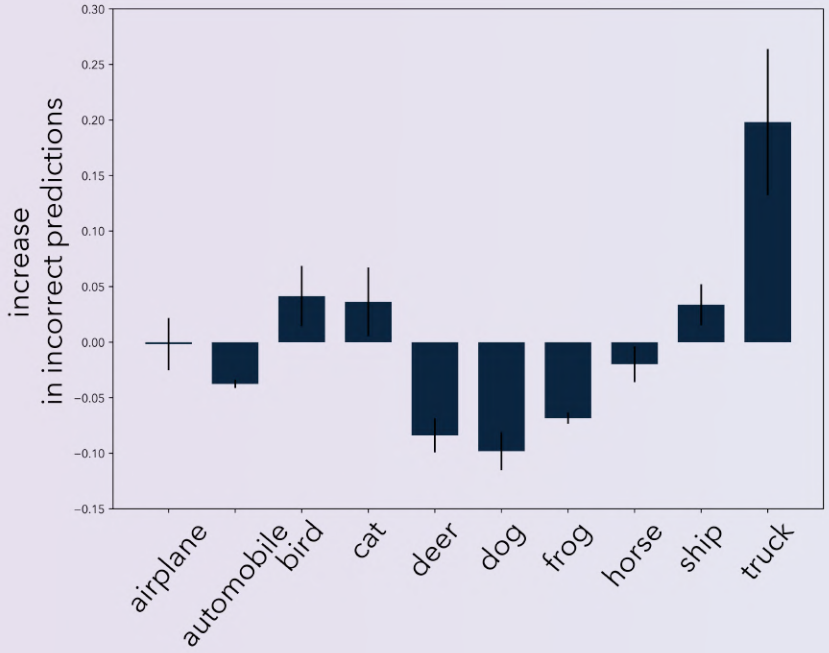
Side-effects: Evaluating

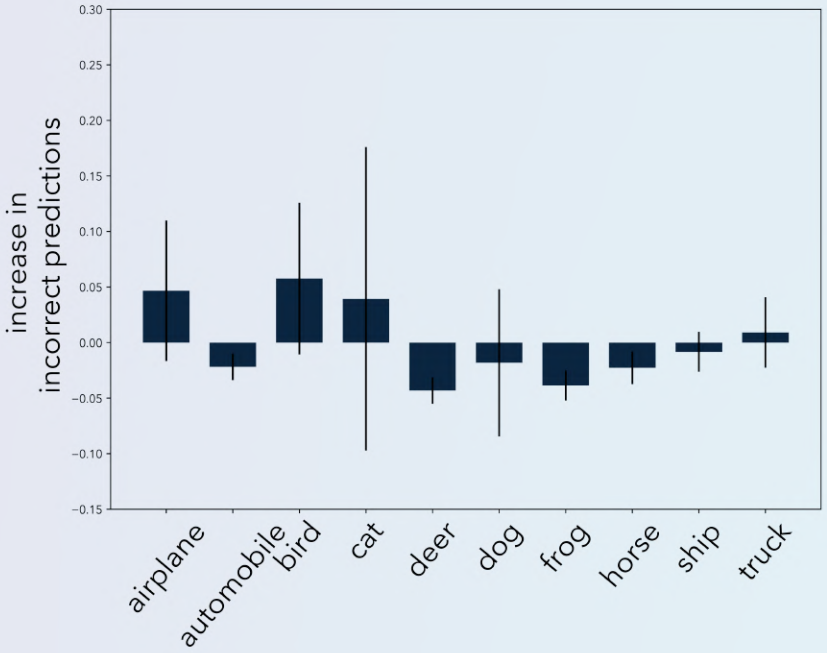
- Are artefacts associated with a class?



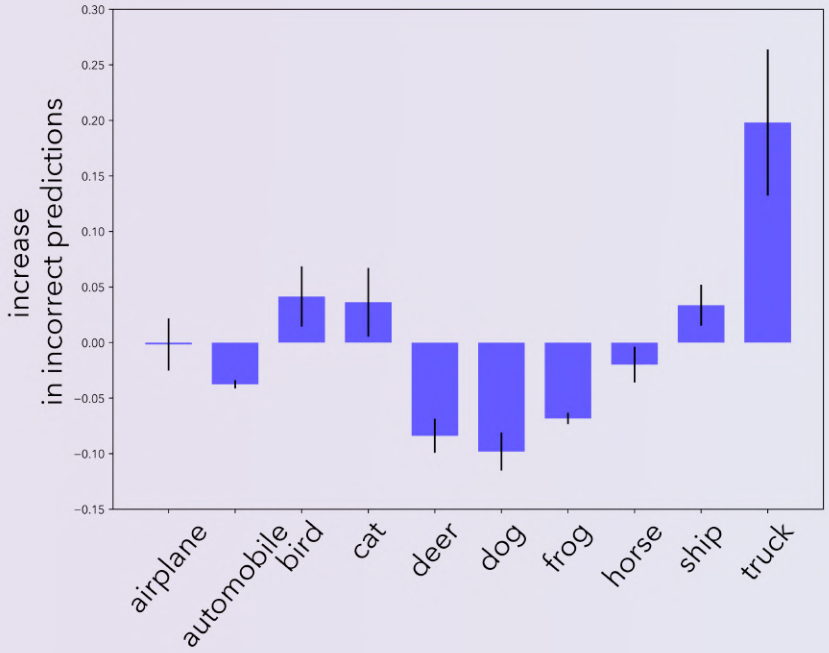


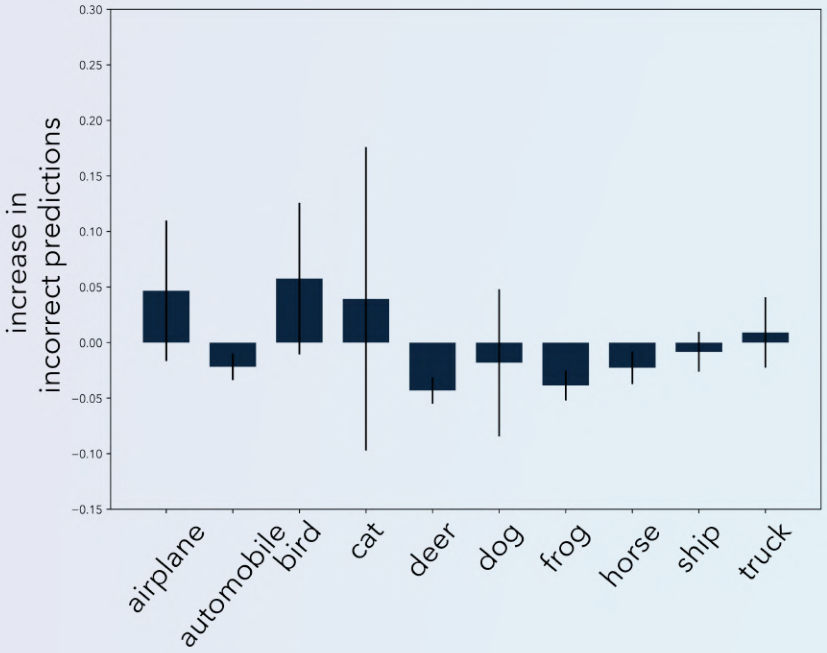
Side-effects: Evaluating



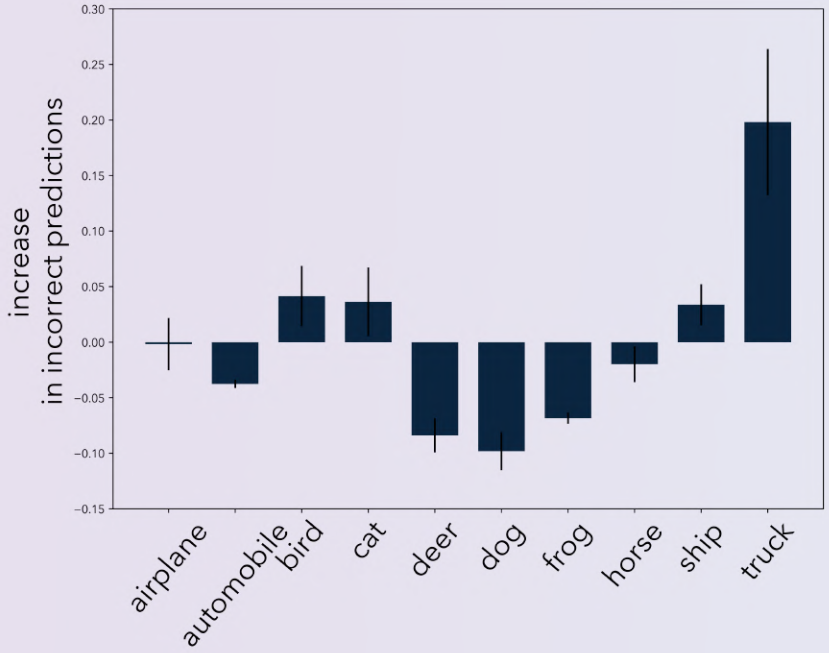


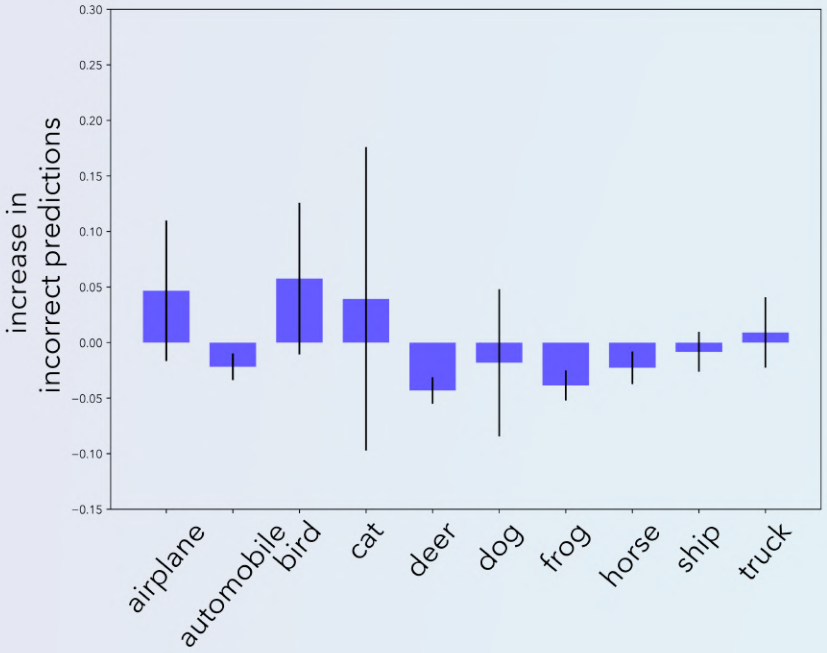
Side-effects: Evaluating





Side-effects: Evaluating





- We create an index for Data Interference

- We create an index for Data Interference

- We propose a new method for measuring occlusion robustness

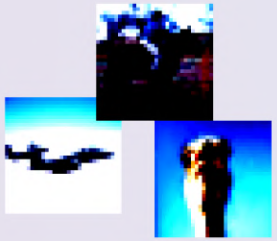
Side-effects: Training

Side-effects: Training

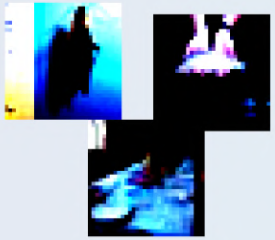
- Distribution gap is not reflective of augmentation quality

Side-effects: Training

CIFAR-10

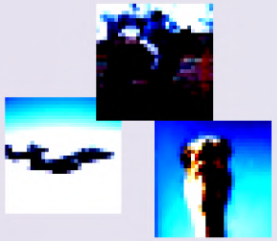


CIFAR-100

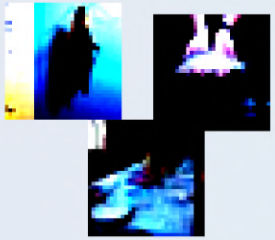


Side-effects: Training

CIFAR-10

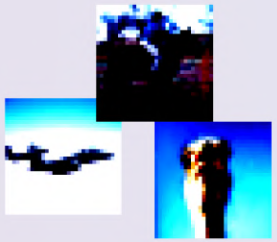


CIFAR-100

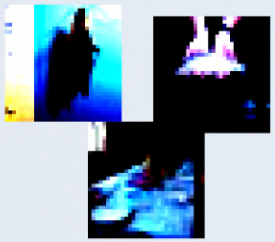


Side-effects: Training

CIFAR-10

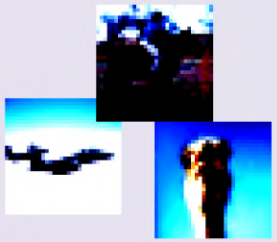


CIFAR-100

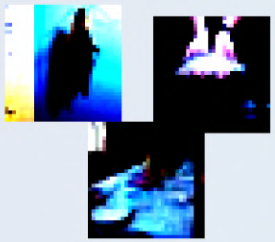


Side-effects: Training

CIFAR-10



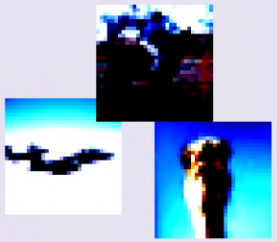
CIFAR-100



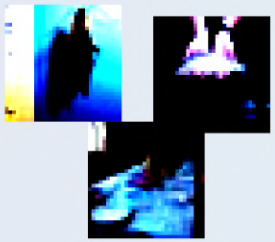
improved generalisation performance

Side-effects: Training

CIFAR-10



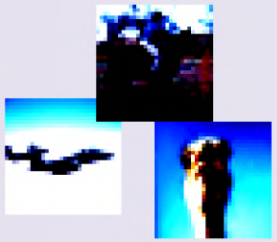
CIFAR-100



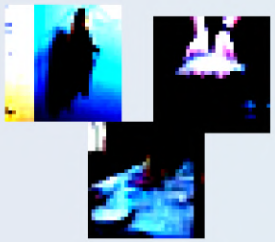
improved generalisation performance

Side-effects: Training

CIFAR-10



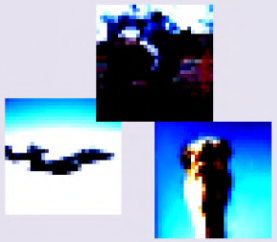
CIFAR-100



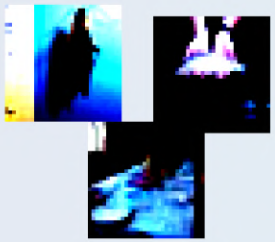
improved generalisation performance

Side-effects: Training

CIFAR-10



CIFAR-100



improved generalisation performance

reduced generalisation performance

Take Away Messages

Take Away Messages

- Incorrect implicit assumptions

Take Away Messages

- Incorrect implicit assumptions
- We explore and address some of them

Take Away Messages

- Incorrect implicit assumptions
- We explore and address some of them
- Distortions can help us understand things better