# Domain-based semi-supervised learning

## Exploiting label invariance in unlabeled data from distributed cameras

Leonardo Taccari

leonardo.taccari@verizonconnect.com

@leotac

**verizon connect**

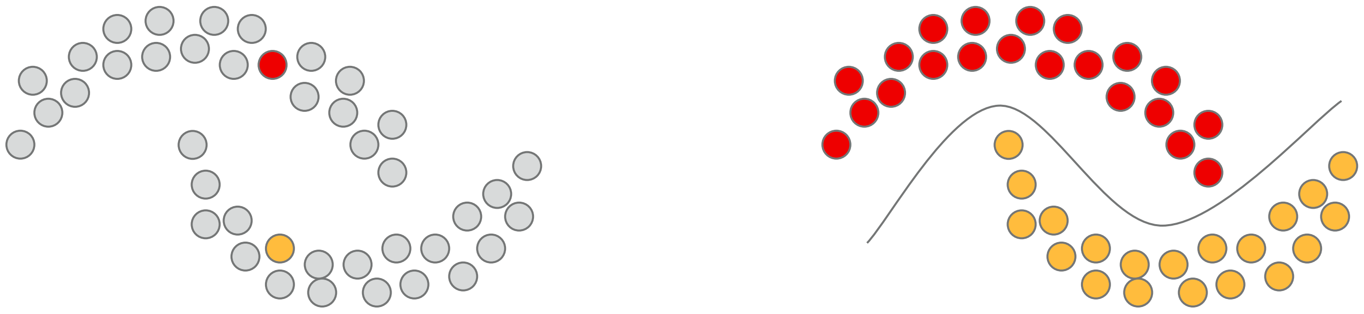
ICCV-IWDSC 2021

# Who I am

- Scientist @ **Verizon Connect**, part of the Verizon group
- *Connected things* (vehicles, assets) for commercial customers (B2B)
- **Millions** of tracked items across the world with GPS and cameras
- Data science/ML group based in **Florence, Italy**

# Semi-supervised learning

- Combine **supervised** and **unsupervised** learning
- Long history (see book by Chapelle), renewed interest in recent years (along with *self-supervised learning*)

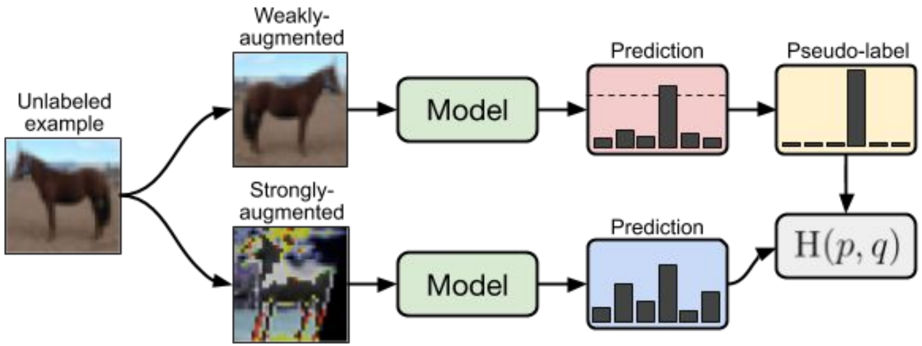


# Recent successful approaches

**Different flavors & ingredients:**

- self-training / pseudolabeling (teacher-student)
- consistency regularization
- aggressive data augmentations

See UDA, MixMatch, FixMatch, Meta Pseudo Labels in the last 2-3 years




FixMatch

# A pragmatic perspective

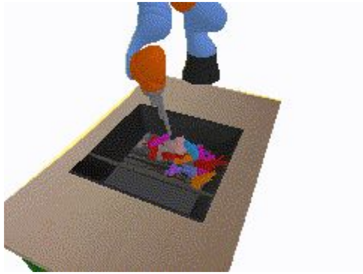
Underlying assumption of SSL work: we know (almost) **nothing** about unlabeled examples

But.. is that true? Often we know a lot! We can **CHEAT** a little bit**.**

💡 Suprisingly often, we can find “**target consistent**” groups in our unlabeled data, where the label is *consistent*, although *unknown.*

# A few examples

Spatially aligned images (same location)



Different point of views of the same scene

Different images from the same device

Similar ideas have also been exploited for **self-supervised learning** approaches (e.g., Time-Contrastive Networks, Geography-Aware Self-supervised Learning)

# DSSL core idea

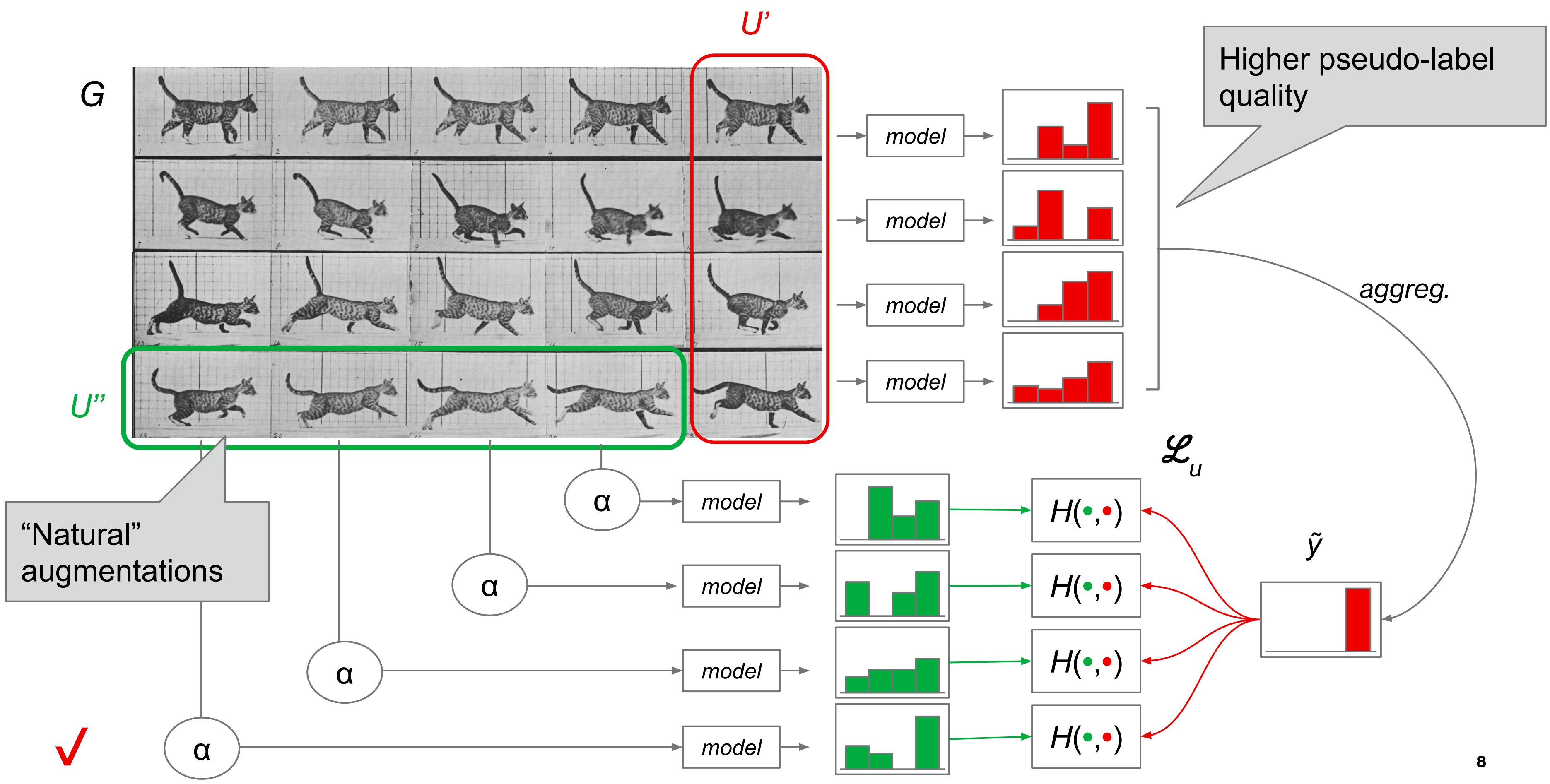
For each batch:

N labeled examples

M target-consistent groups of unlabeled examples

Compute:

$$\mathcal{L} = \underbrace{\mathcal{L}_s}_{\text{standard supervised loss}} + \underbrace{\lambda_u \mathcal{L}_u}_{\text{unsupervised loss term (consistency) exploiting known invariants}}$$



U'
G
U''
model
model
model
model
Higher pseudo-label quality
aggreg.
$\mathcal{L}_u$
α
α
α
α
model
model
model
model
H(•,•)
H(•,•)
H(•,•)
H(•,•)
ỹ
"Natural" augmentations

# Experiments

3 practical cases on distributed cameras + 1 bonus

# Weather classification

**Task:** classify weather in dashcam images

Thousands of unlabeled short videos

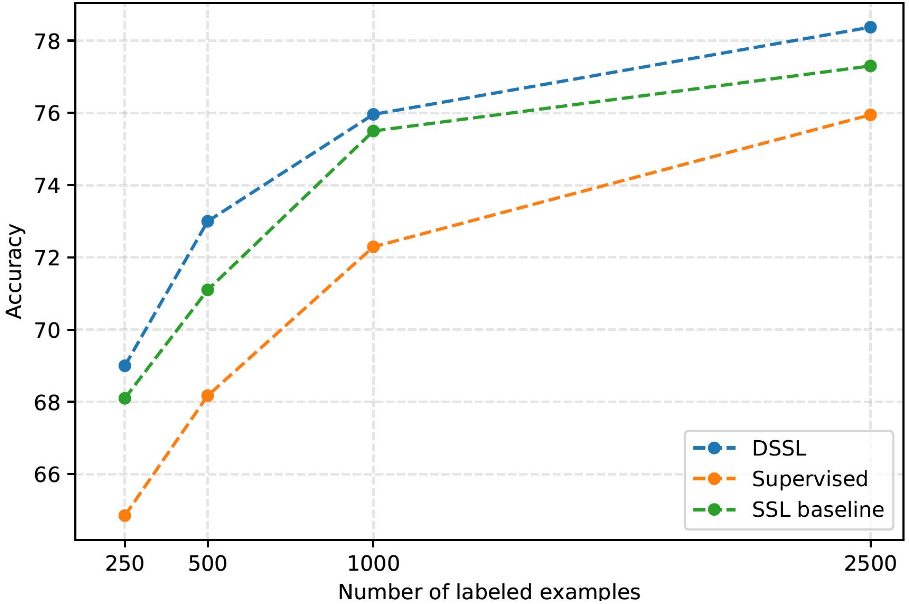
**Weather doesn’t change locally:** frames from the same video are “target-consistent”.

Data from BDD100k

Berkeley DeepDrive

# Weather classification

SSL baseline ≈ FixMatch "with no bells and whistles"




DSSL (kind of) works even with no data augmentation at all

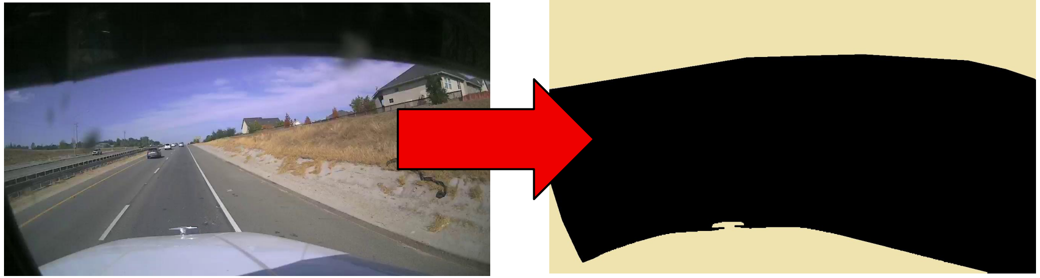
| | $B_u$ | $K', K''$ | Augmentation | Accuracy |
|---|---|---|---|---|
| DSSL | 8 | (4, 4) | RandAugment | **73.0** |
| SSL baseline | 32 | - | RandAugment | 71.1 |
| Supervised | - | - | RandAugment | 68.2 |
| DSSL | 8 | (4, 4) | Weak | **70.9** (-2.1) |
| SSL baseline | 32 | - | Weak | 68.2 **(-2.9)** |
| Supervised | - | - | Weak | 66.1 (-2.1) |
| DSSL | 8 | (4, 4) | None | **70.0** (-3.0) |
| SSL baseline | 32 | - | None | 63.8 **(-7.3)** |
| Supervised | - | - | None | 64.6 (-3.6) |

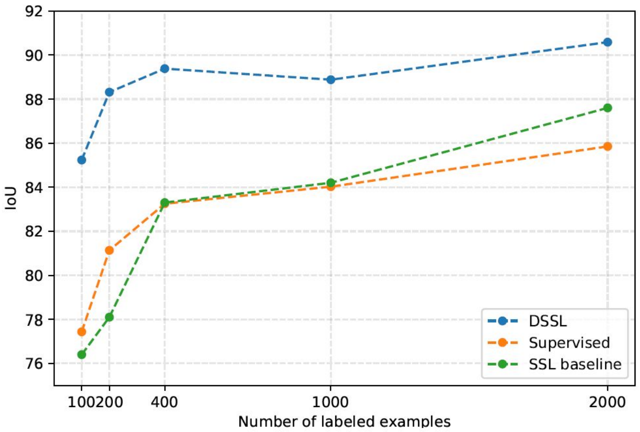
# Ego-vehicle segmentation

**Task:** segment *ego-vehicle* in dashcam images

Millions of unlabeled images (internal dataset)

**Camera position doesn't change:** frames from the same device are "target-consistent".





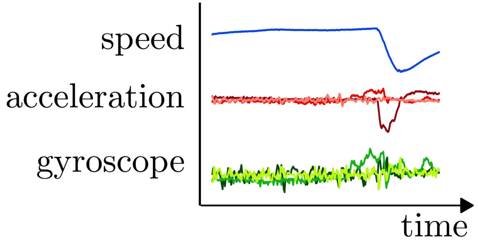
# Vehicle classification

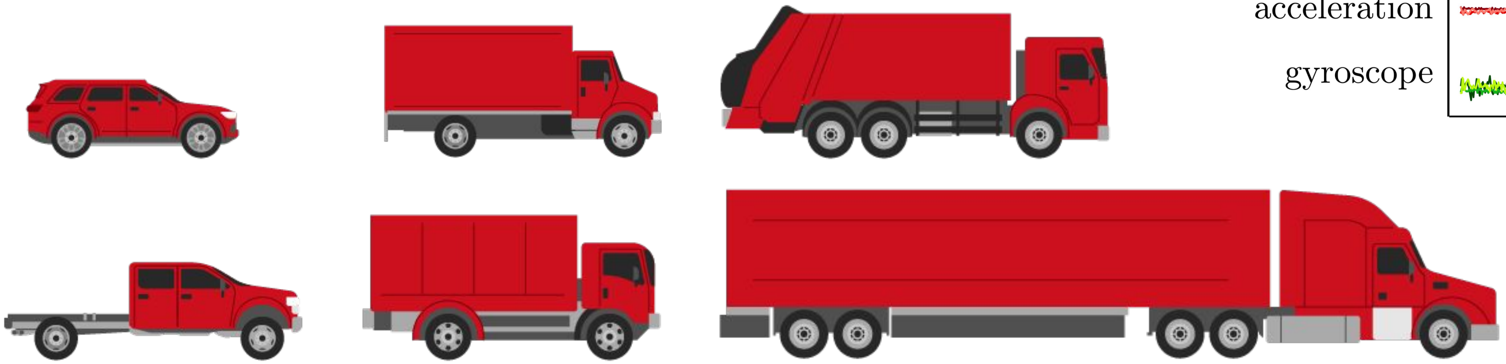
Streams of IMU + GPS data from connected vehicles

**Task:** what kind of vehicle generated the data? (10 classes)

For a subset of vehicles, we know make + model.
Labeling *more* data is not possible!

**Vehicle type doesn't change**: data collected from the same device over time are "target-consistent".



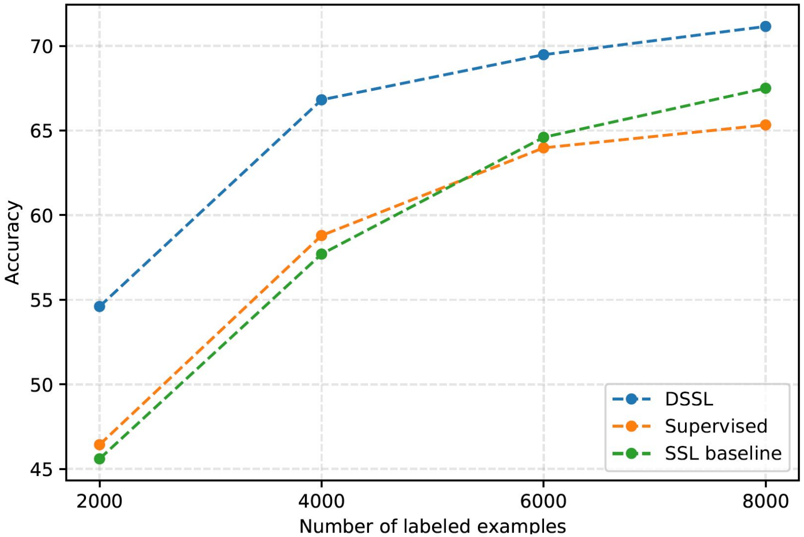
*Images from https://www.fleetequipmentmag.com/making-sense-of-truck-classification/*

# Vehicle classification

These are not images: no **standard recipe** for data augmentations

We used strong augmentations for IMU+GPS data that worked well for other tasks (even self-supervised learning)

They're sufficient for DSSL, apparently not for the SSL baseline



# Bonus: CIFAR10 & CIFAR100 (proof of concept)

Standard setup: only use a subset as “labeled”

Assume an **oracle** can tell us, for each batch of unlabeled data, which samples belong to the same class (we’re obviously cheating!)

Only **DSSL** uses the extra information

| | CIFAR10 | | CIFAR100 | |
|---|---|---|---|---|
| | 250 labels | 4000 labels | 2500 labels | 10000 labels |
| DSSL * | **96.75**±0.28 | **96.80**±0.03 | **76.50**±0.02 | **76.89**±0.01 |
| FixMatch (RA) *† | 94.30±0.45 | 95.75±0.05 | 66.51±0.32 | 72.41±0.13 |
| Meta Pseudo Labels† | - | 96.11±0.07 | - | - |
| FixMatch (CTA) † | 94.93±0.33 | 95.69±0.15 | 71.36±0.24 | 76.82±0.11 |
| MixMatch † | 88.95±0.86 | 93.58±0.10 | 60.06±0.37 | 71.69±0.33 |
| UDA † | 91.18±1.08 | 95.12±0.18 | 66.87±0.22 | 75.5±0.25 |

Same implementation

Published results

**DSSL is a simple yet effective strategy to exploit domain knowledge in SSL**

# Domain-based semi-supervised learning

## Exploiting label invariance in unlabeled data from distributed cameras

Leonardo Taccari

leonardo.taccari@verizonconnect.com

@leotac

verizon connect

ICCV-IWDSC 2021