



DINOv2

LLVM Paper Discussion

10.30.23



Academic Researcher

Alexandre Kaiser amk1004

Innovations proposed by the paper

Automatic data curation process

- Curated datasets (ImageNet, etc...)
- 1.2 billion raw unfiltered web images
- Deduplication (Pizzi et al. 2022)
- Self-supervised image retrieval
- Faiss library (Johnson et al. 2019)
- GPU accelerated indices (Jegou et al. 2010)

Aggregation of prior work leading to SOTA performance

- DINO loss
- iBOT loss (Zhou et al. 2021)
- SwAV centering (Caron et al. 2020)
- KoLeo regularizer (Sablayrolles et al. 2018)
- Adapting resolution (Touvron et al. 2019)

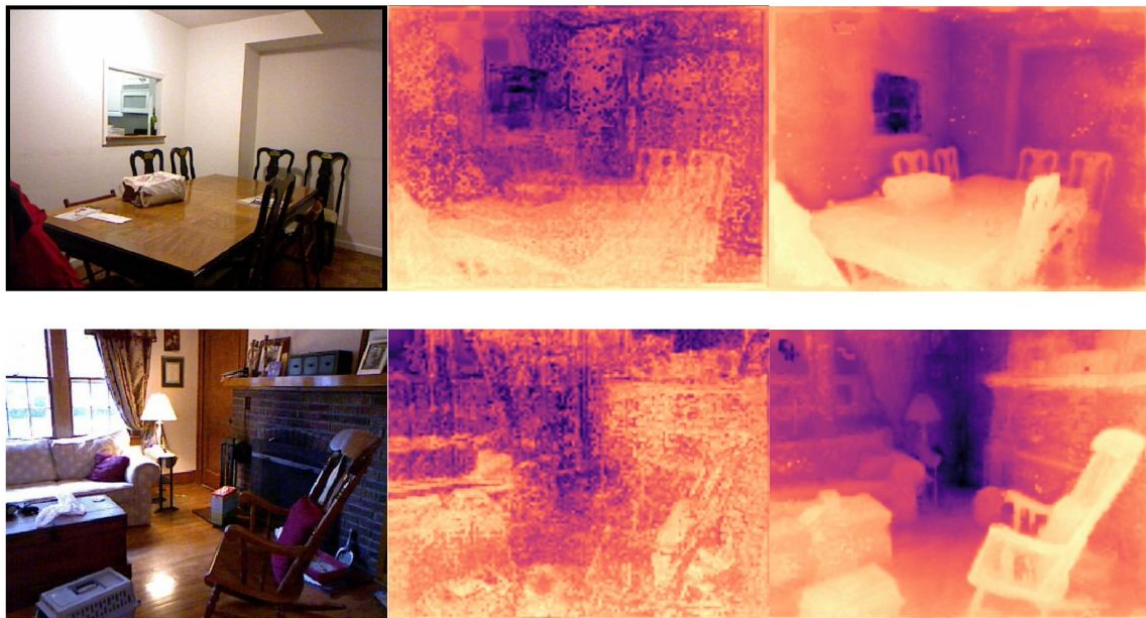
What are the features capable of?

- Segmentation (w/ labels)
- Depth estimation
- Patch matching
- PCA components aligned and ordered to human expectation
- Compatible with linear classifiers



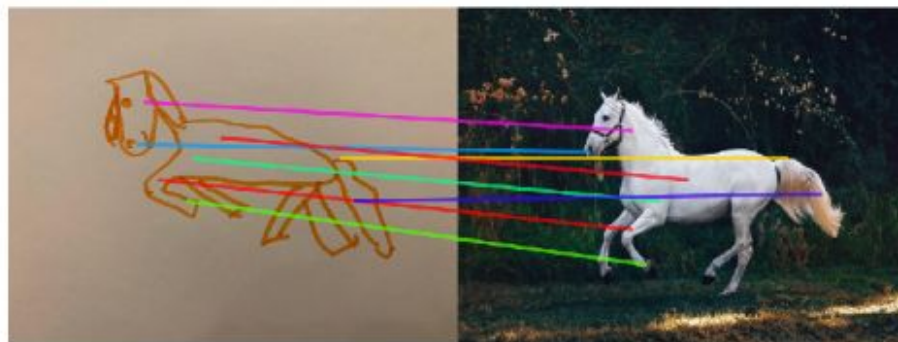
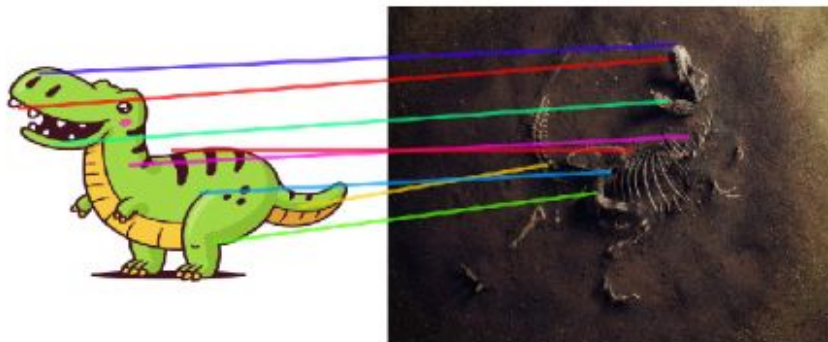
What are the features capable of?

- Segmentation (w/ labels)
- Depth estimation
- Patch matching
- PCA components aligned and ordered to human expectation
- Compatible with linear classifiers



What are the features capable of?

- Segmentation (w/ labels)
- Depth estimation
- Patch matching
- PCA components aligned and ordered to human expectation
- Compatible with linear classifiers



What are the features capable of?

- Segmentation (w/ labels)
- Depth estimation
- Patch matching
- PCA components aligned and ordered to human expectation
- Compatible with linear classifiers



What are the features capable of?

- Segmentation (w/ labels)
- Depth estimation
- Patch matching
- PCA components aligned and ordered to human expectation
- Compatible with linear classifiers

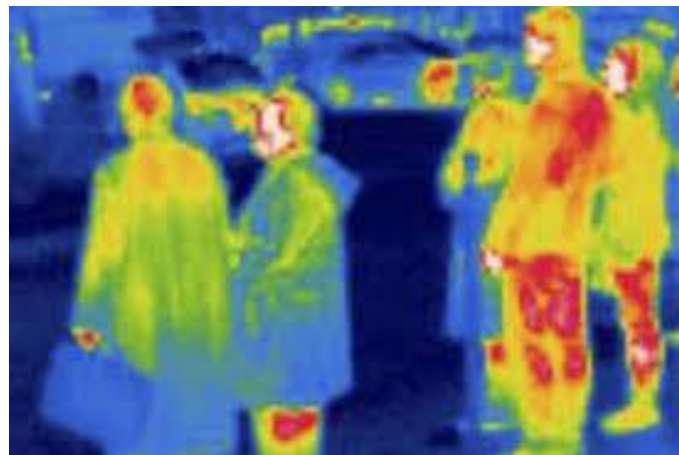


meaning the underlying information is *readily available*.

Future directions

- SOTA for other modalities with large uncured public datasets
 - Audio
 - Video*
 - Brain
- Unified representations (image+text)
- Infer other modalities
 - Heat
 - IMU
 - Colorization
 - Light transport
 - Emotion

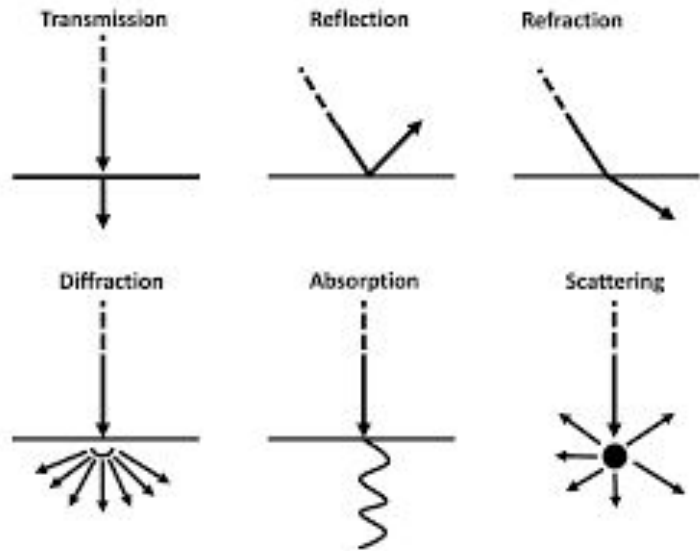
Heat (inspired by ImageBind)



Colorization



Inverse light transport



Softness



Material Properties

Chemical

- Corrosive
- Reactivity
- pH

Physical

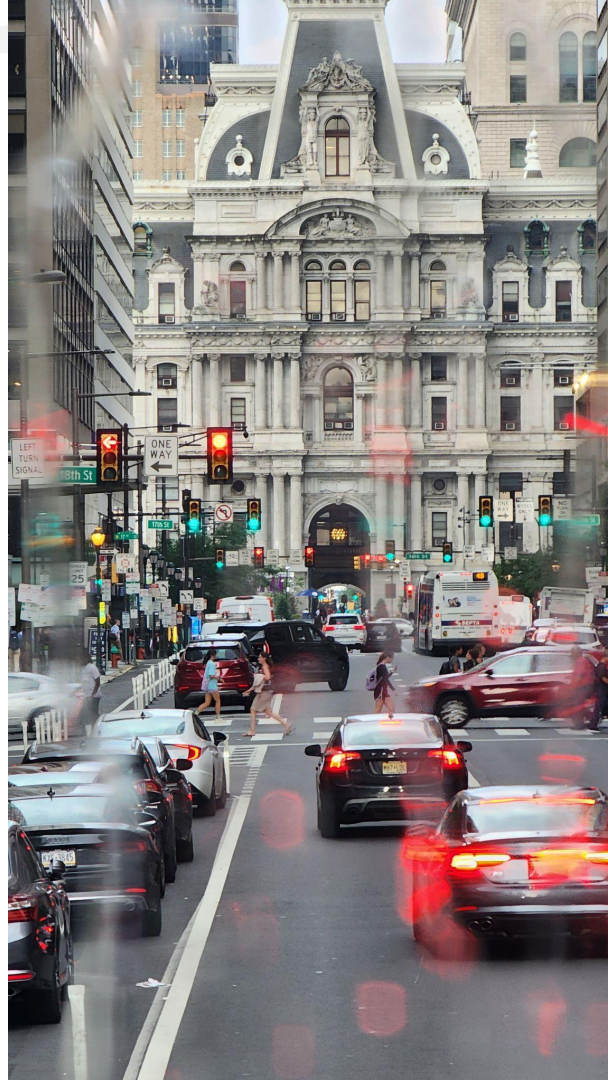
- Density
- Conductivity
- Optical
- Acoustic
- Combustibility
- Tension

Mechanical

- Elasticity
- Plasticity
- Yield
- Ductility
- Hardness
- Toughness

Issues with these mappings

Strange artifacts
can act like
non-gaussian
noise



What about more abstract concepts?

Emotion

Habitability

Style

Beauty

Realism

