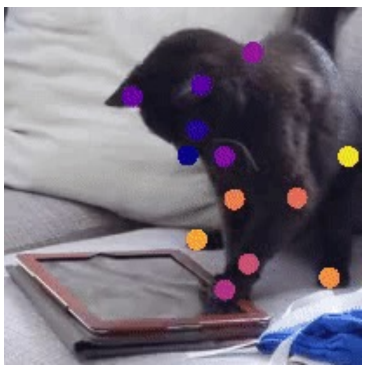
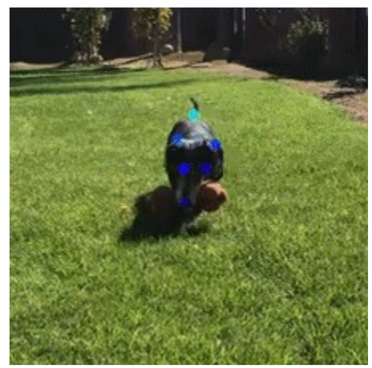
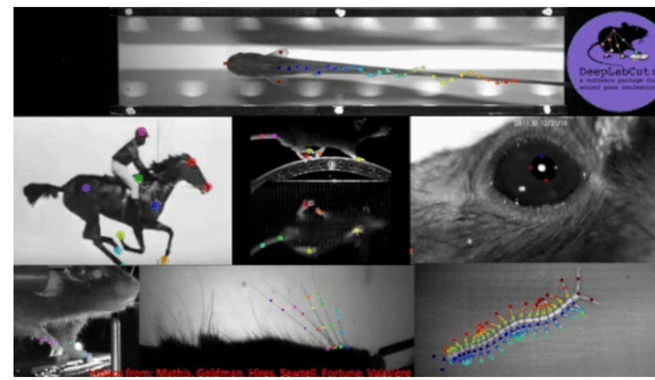
Looking at animals in 3D

Silvia Zuffi
IMATI-CNR
28 Sep 2022

Looking at animals

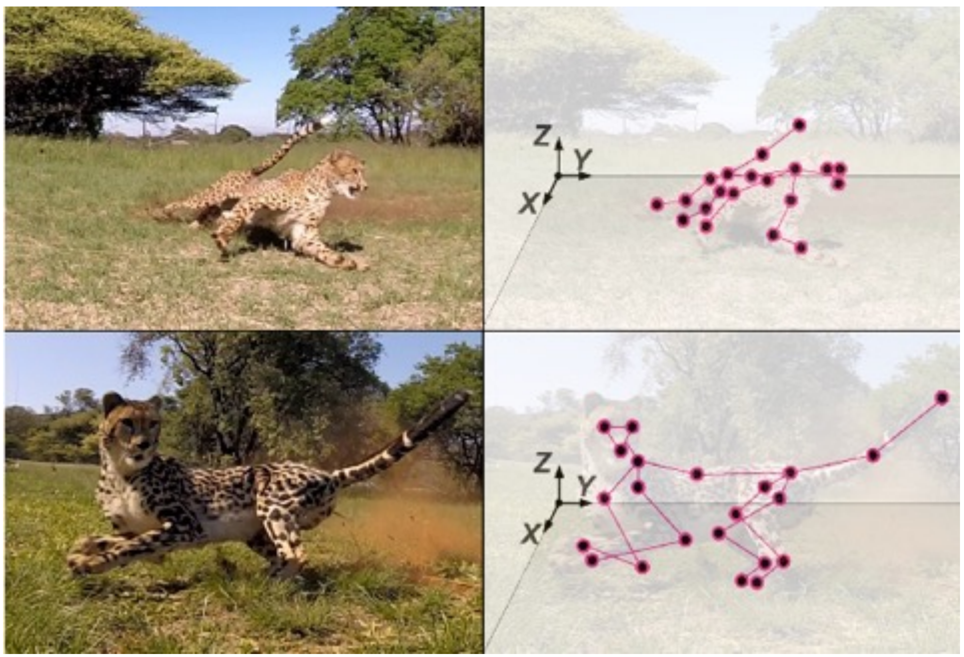
- 2D pose estimation dominates animal behavior research



Looking at animals in 3D

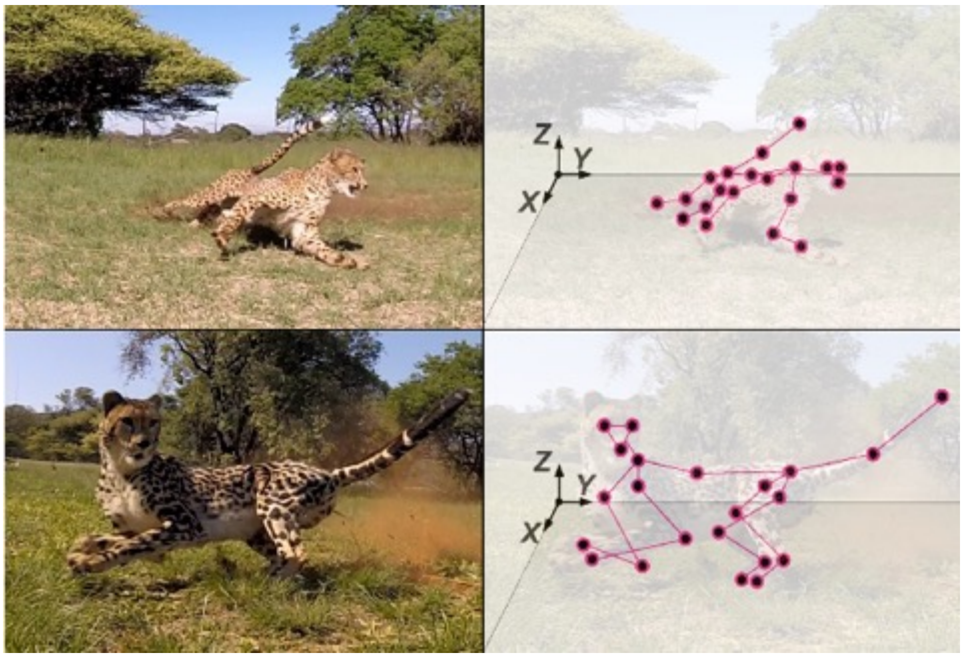
- Estimate the 3D location of the body joints



D. Joskra et al., AcinoSet: A 3D Pose Estimation Dataset and Baseline Models for Cheetahs in the Wild, ICRA 2021

Looking at animals in 3D

- Estimate the 3D location of the body joints



D. Joskra et al., AcinoSet: A 3D Pose Estimation Dataset and Baseline Models for Cheetahs in the Wild, ICRA 2021

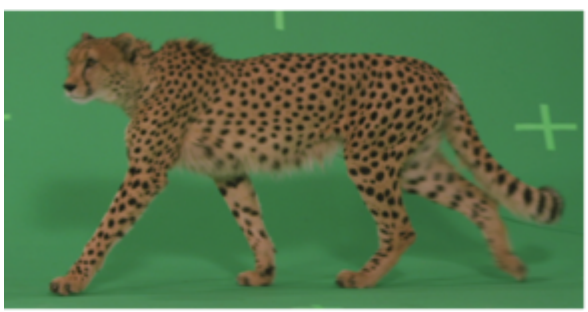
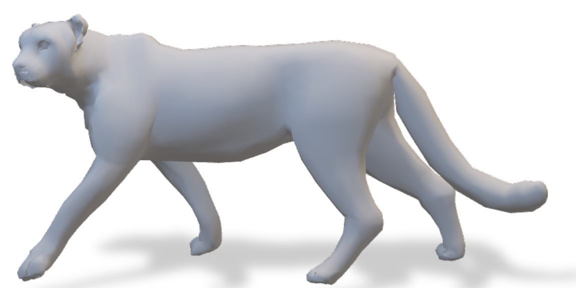
Not yet sufficient!

- Couples 3D pose with body shape
- Hard to lift the 2D landmarks to 3D in monocular settings

Looking at animals in 3D

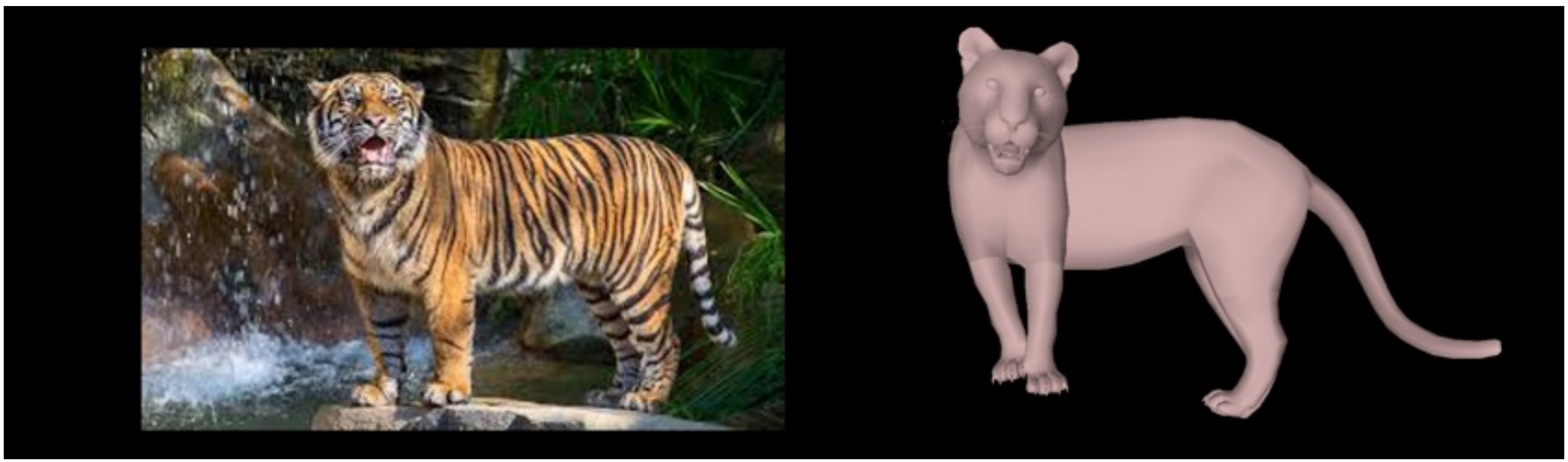
With a model-based approach

- Estimate 3D pose using a parametric model
- The model represents **prior knowledge** about the body shape and helps in monocular **reconstruction**


Looking at animals in 3D

- Shape is functional to 3D pose estimation from monocular data



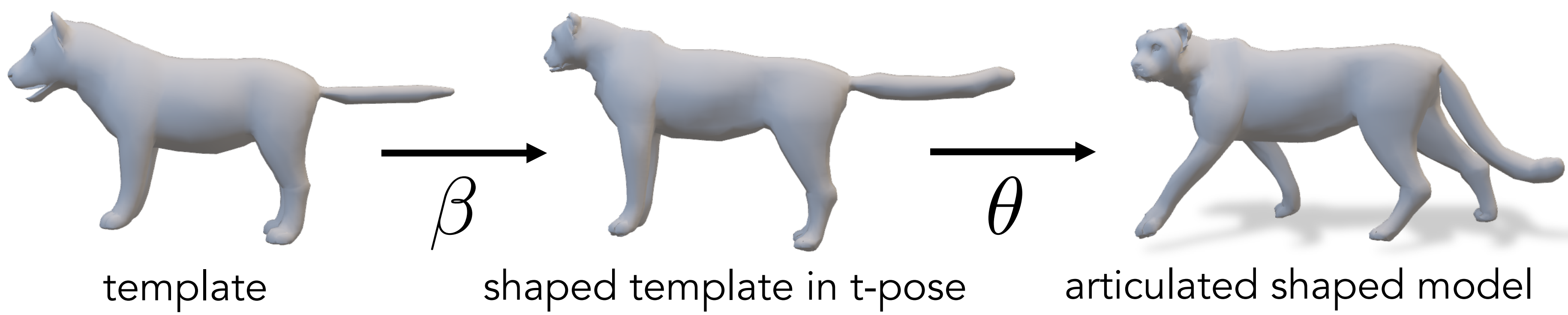
Looking at animals in 3D

- If shape is not correct, then 3D pose is wrong: we need to predict accurate shape even if we are only interested in 3D pose!

A model-based approach

- What is a 3D model of an animal?
  - A mathematical formulation that, given disentangled shape and pose parameters, deforms a template to return a 3D object

$$\mathbf{v} = \mathcal{M}(\beta, \theta)$$

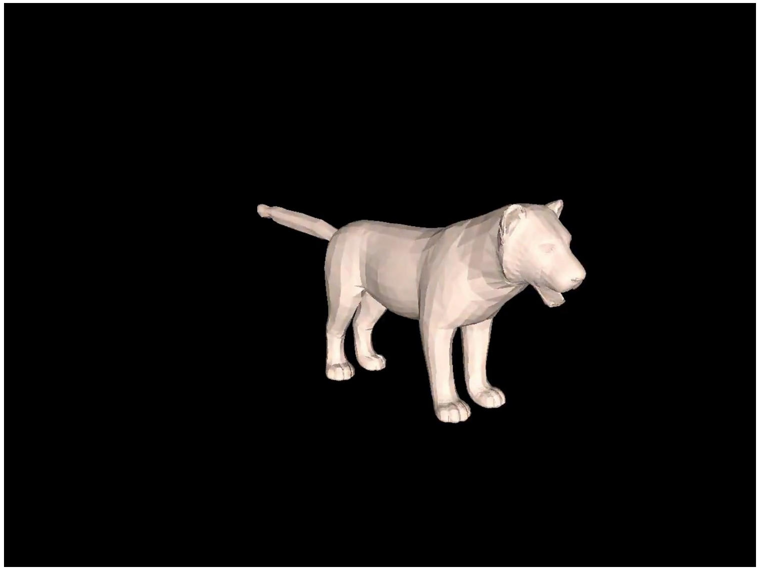


Model-based 3D posture analysis

$$\mathbf{v} = \mathcal{M}(\beta, \theta)$$

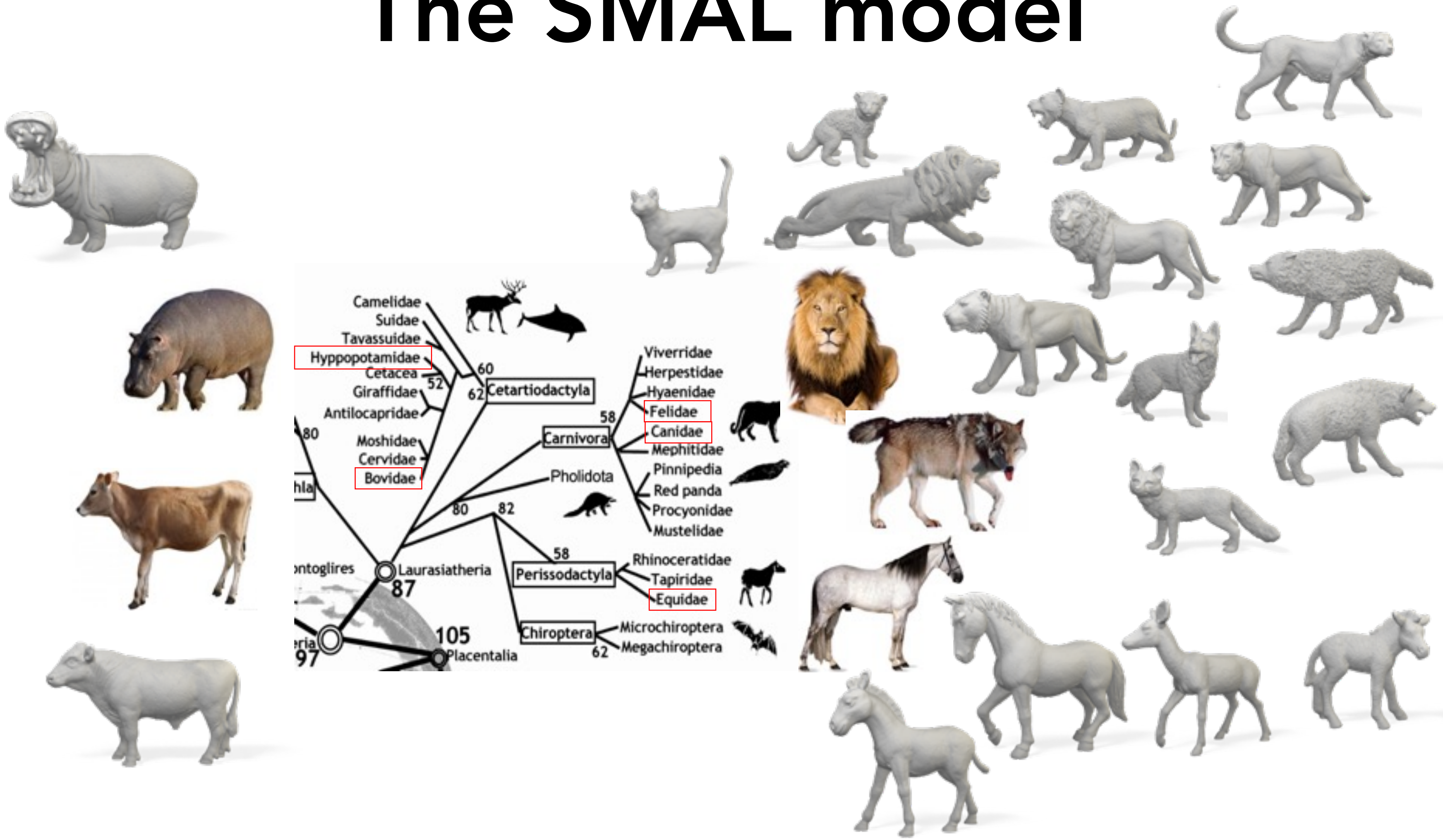
- With the disentagled representation we can compare 3D pose of different species, provided they are represented with the same skeleton

The SMAL model

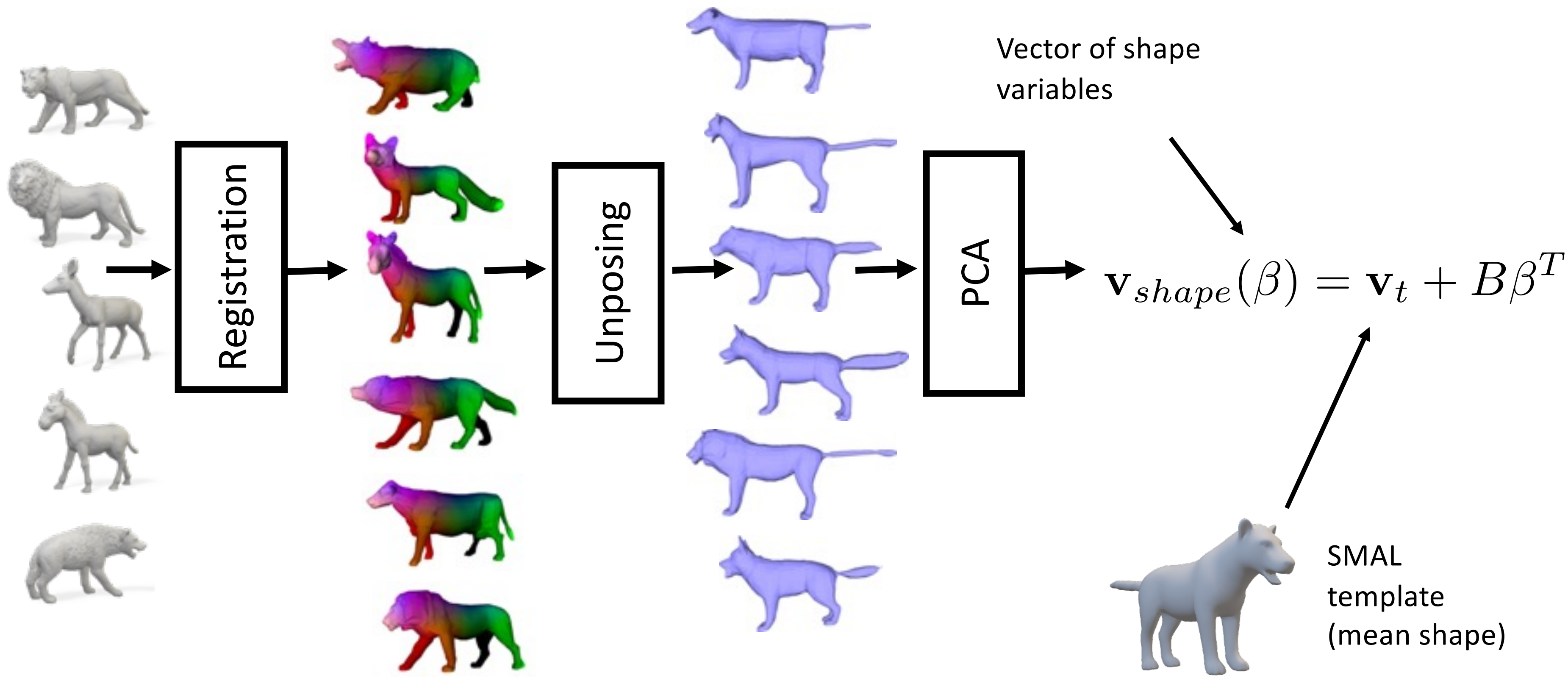


S. Zuffi, A. Kanazawa, D. Jacobs, M.J. Black, 3D Menagerie: Modeling the 3D Shape and Pose of Animals, CVPR 2017

The SMAL model



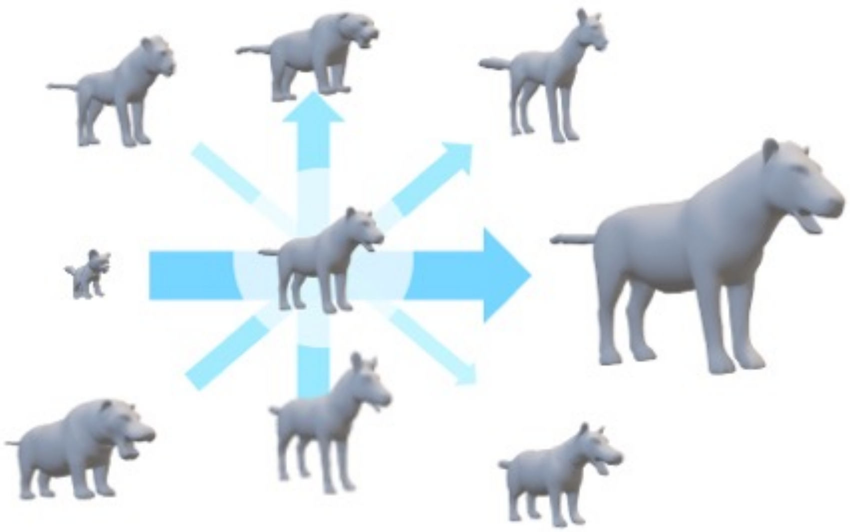
Learning SMAL



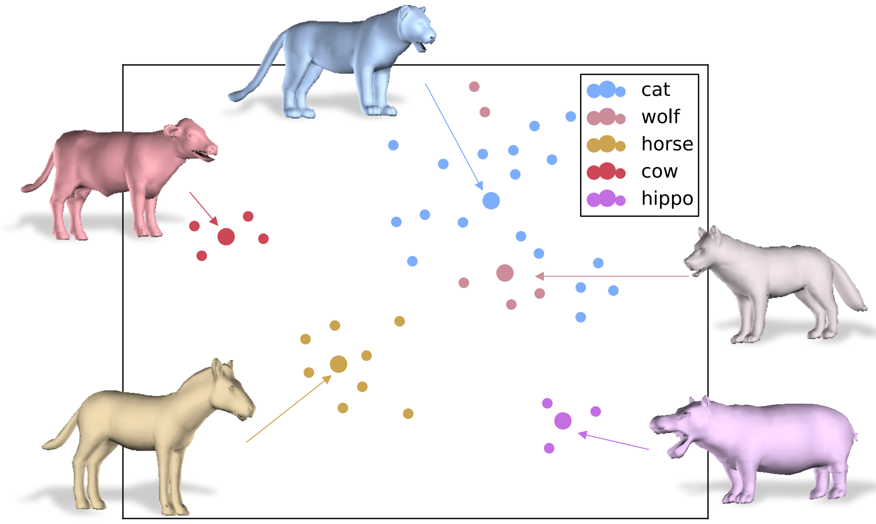
Shape space

$$\mathbf{v}_{shape}(\beta) = \mathbf{v}_{template} + B_s \beta^T$$

Shape space first 4
principal components

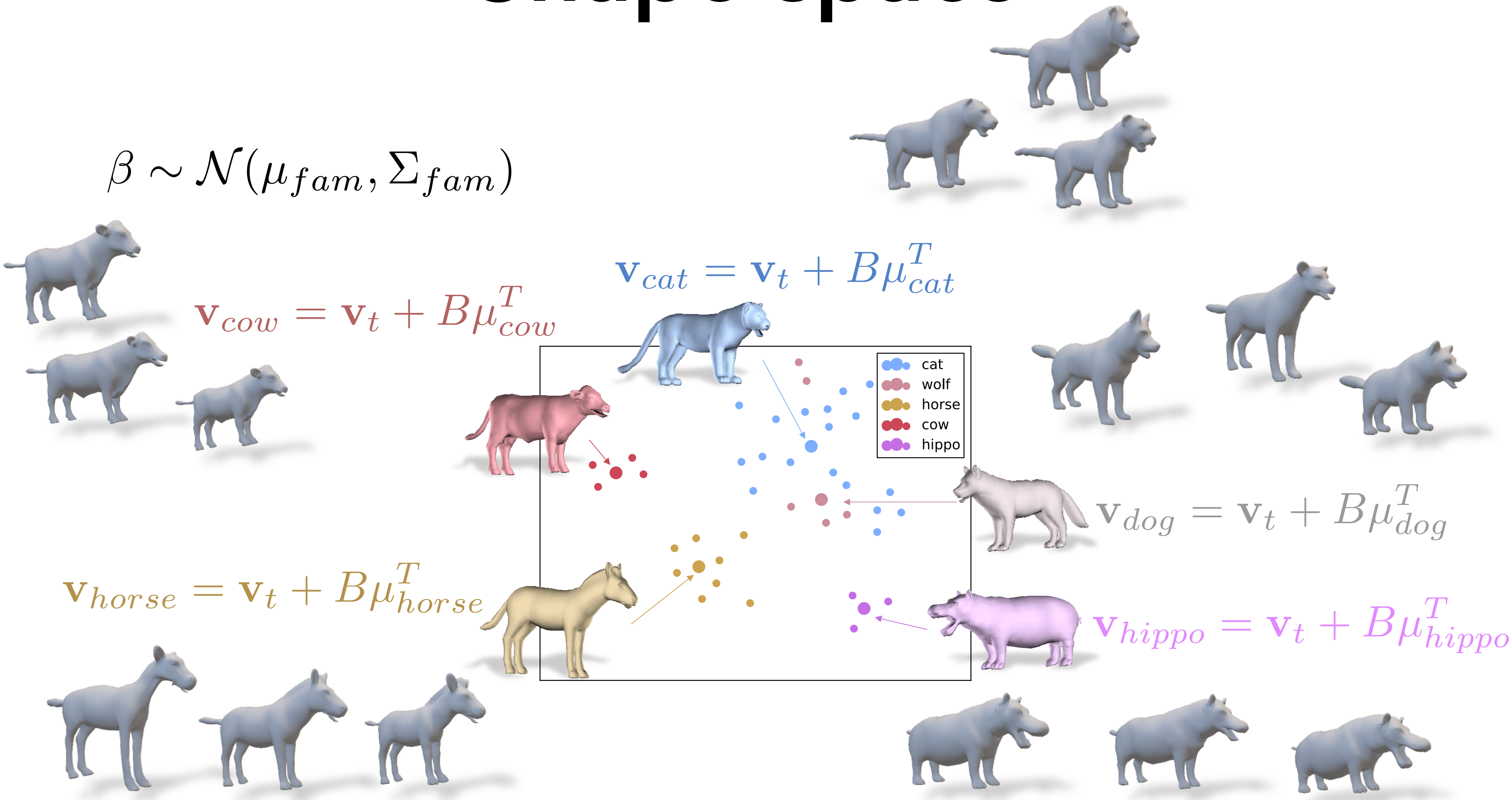


2D t-sne plot of the shape
variables for the training samples



Shape space

$$\beta \sim \mathcal{N}(\mu_{fam}, \Sigma_{fam})$$



Model fit to images

SMALR

- Optimization
- Multiple images

SMALST

- Grevy's zebra
- Limited appearance variation
- Trained with 3D data

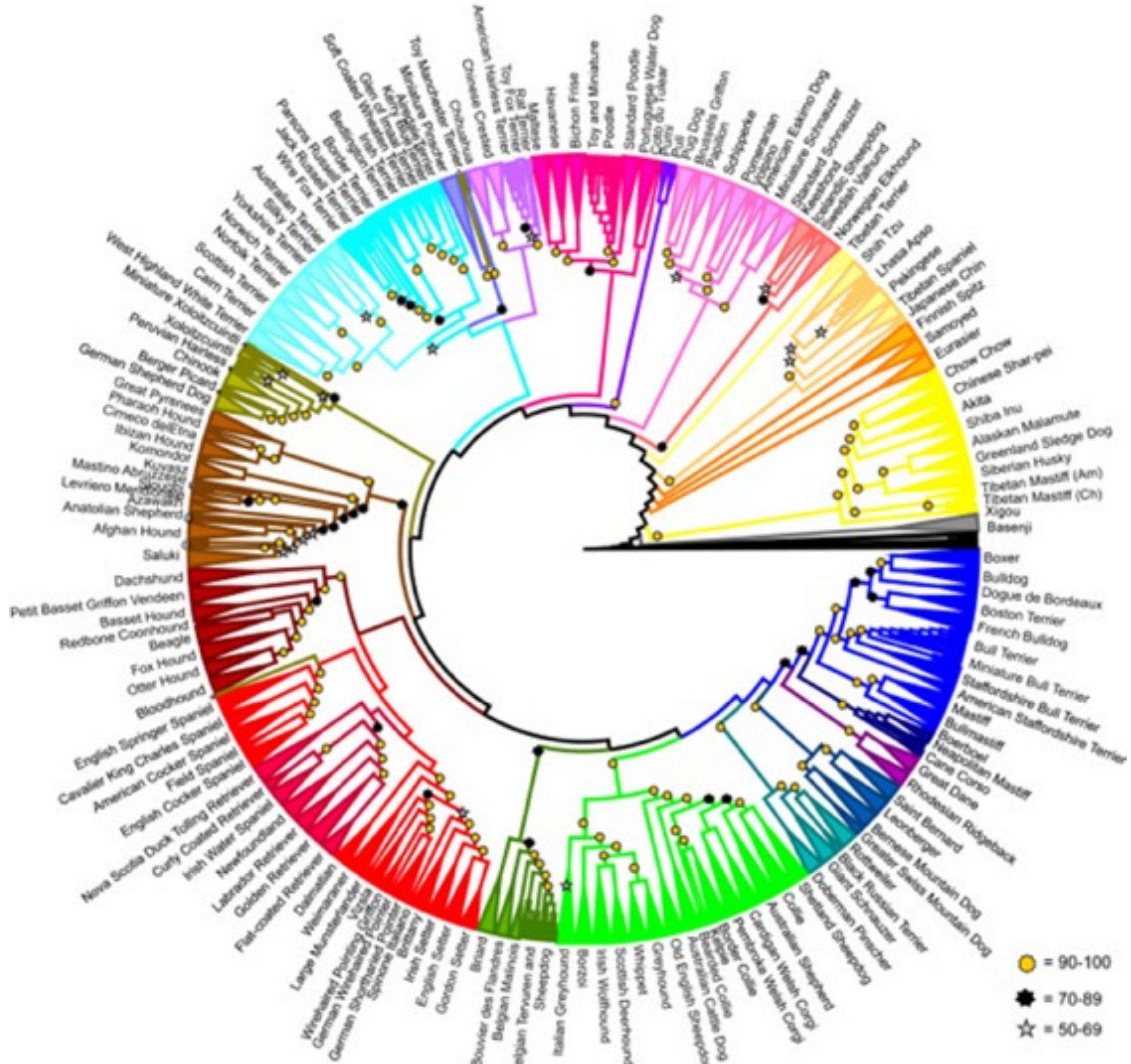
BARC

- Dogs
- High appearance variation
- Trained with 2D data

BARC: Learning to Regress 3D Dog Shape from Images by Exploiting Breed Information

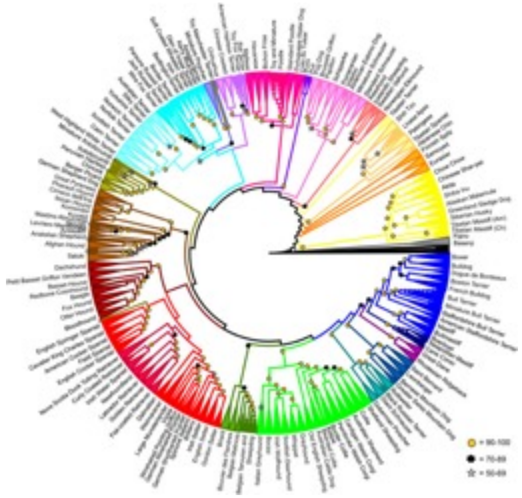
N. Ruegg, S. Zuffi. K. Schindler, M.J. Black, BARC : Learning to Regress 3D Dog Shape from Images by Exploiting Breed Information, CVPR 2022

BARC



H. G. Parker et al, Genomic analyses reveal the influence of geographic origin, migration, and hybridization on modern dog breed development. Cell Reports, 4(19):697–708, 2017

BARC



EuropeanMastiff
Alpine
Drover
UKRural
UKRural
Continental Herder
PointSetter
Retriever
Asian Spitz
Basenji
Eurasier
Samoyed
Finnish Spitz
Spaniel
Basenji
Asian Toy
Tibetan Terrier
NordicSpitz
Schnauzer
SmallSpitz
ToySpitz
Hungarian
Poodle
AmericanToy
AmericanTerrier
Pinscher
Terrier
NewWorld
Mediterranean
ScentHound

Side information

Husky

Similar shape

Husky

Different shape

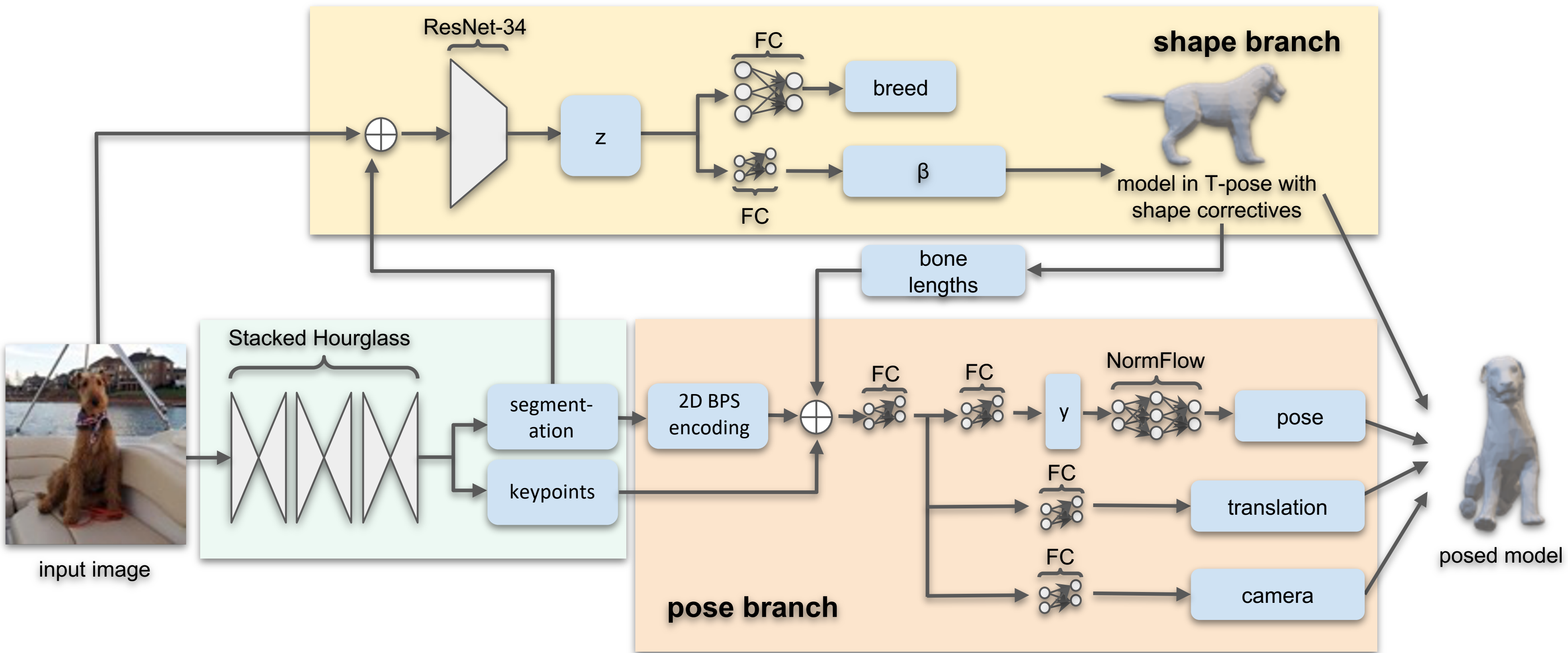
French bulldog

BARC

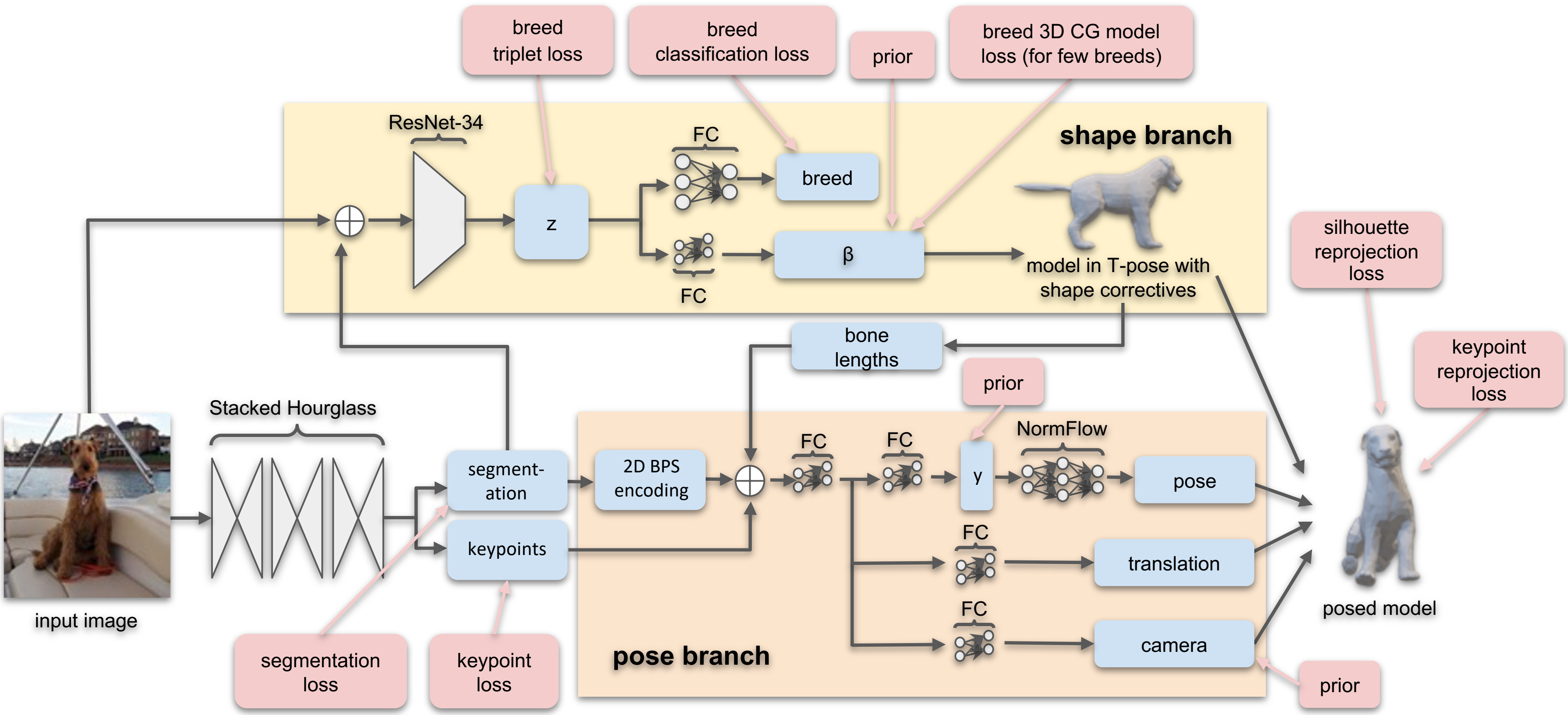
StanfordExtra Dataset

Khosla et al., Novel dataset for Fine-Grained Image Categorization, CVPR 2011
B. Biggs et al., Who left the dogs out? 3D animal reconstruction with expectation maximization in the loop, ECCV 2020

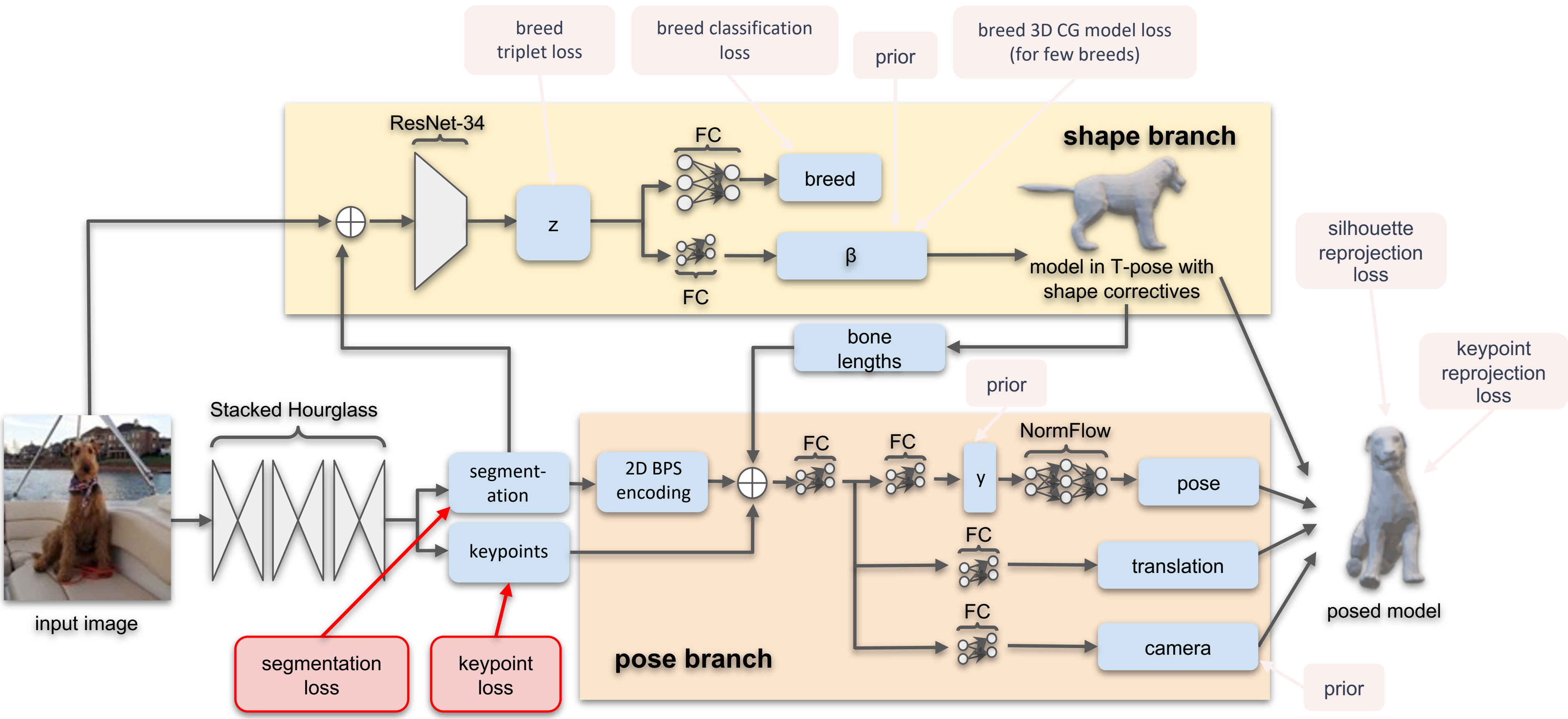
Breed-aware reconstruction



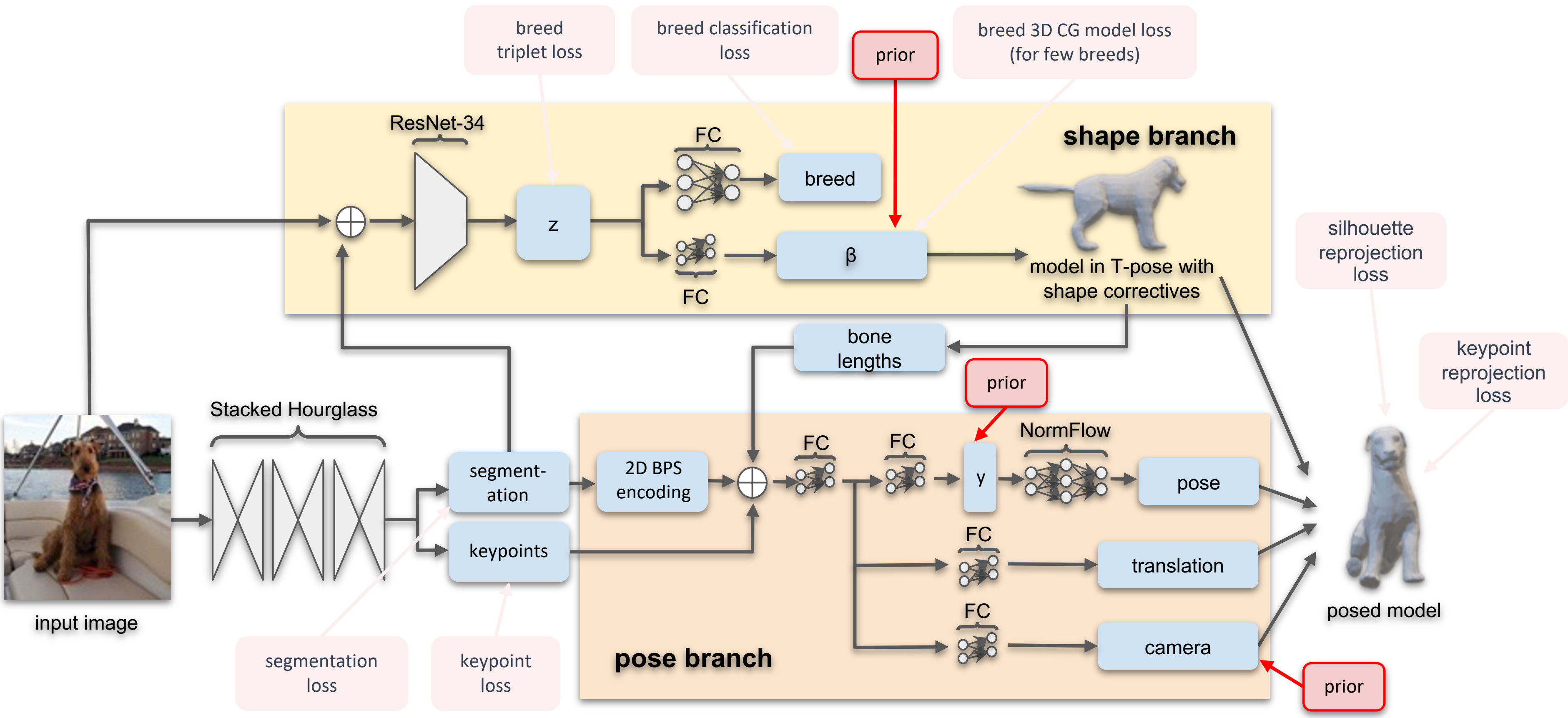
Breed-aware reconstruction



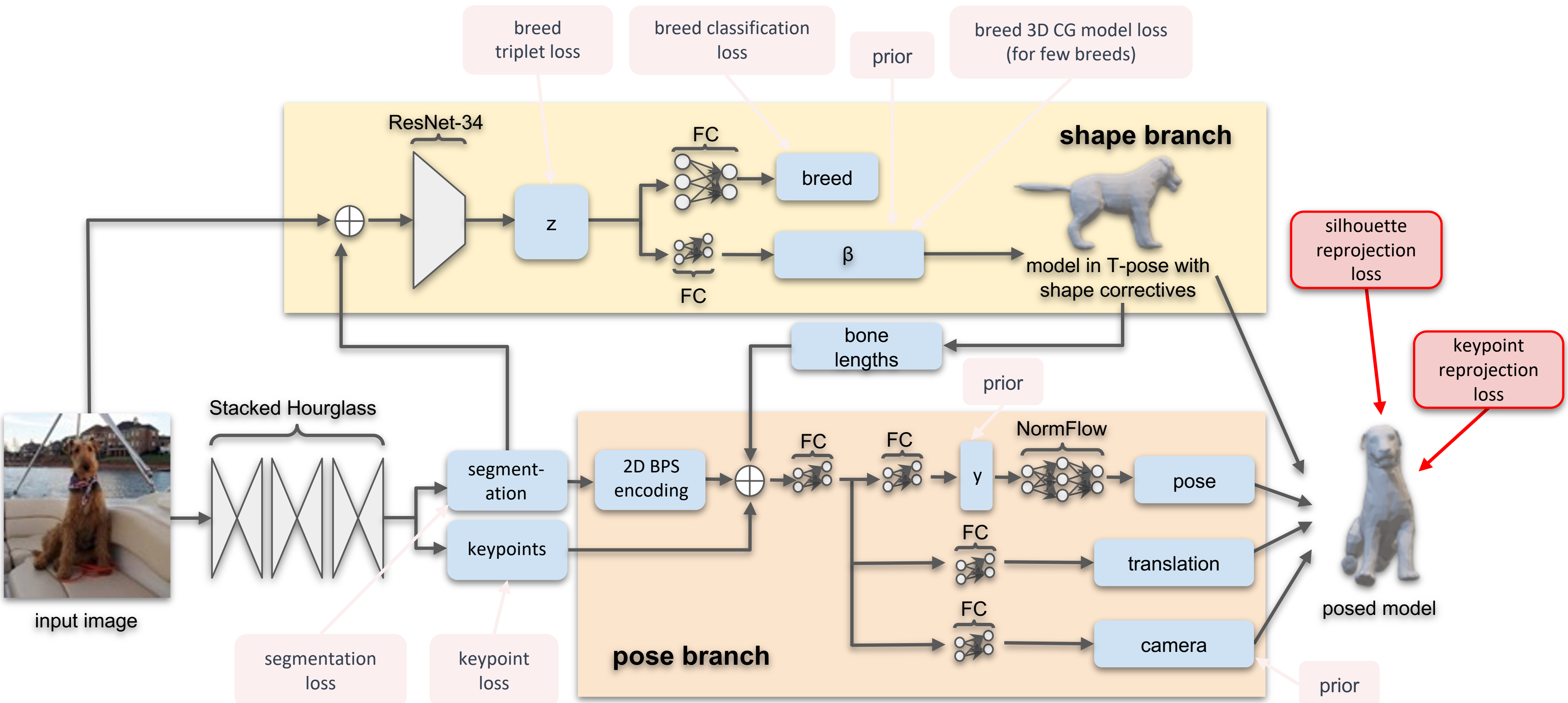
Breed-aware reconstruction



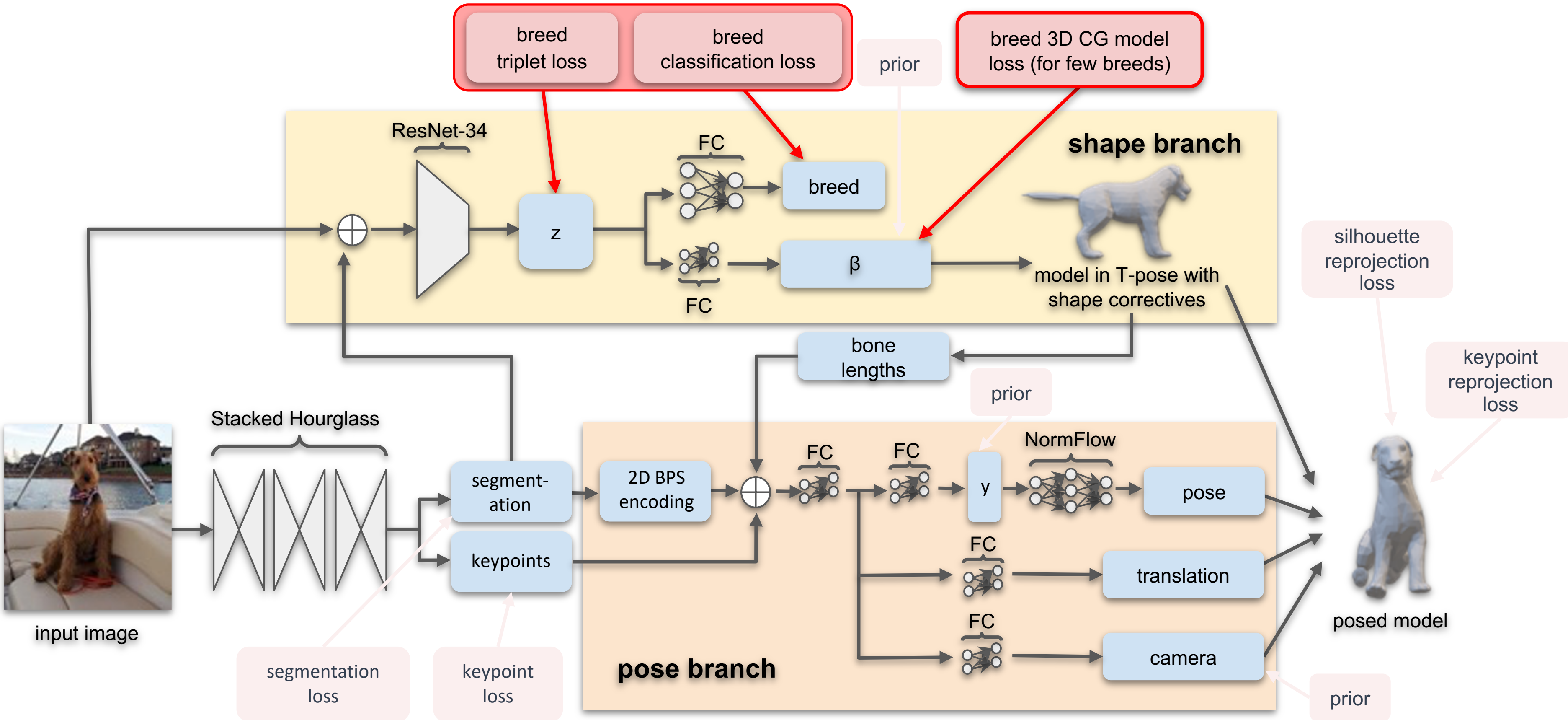
Breed-aware reconstruction



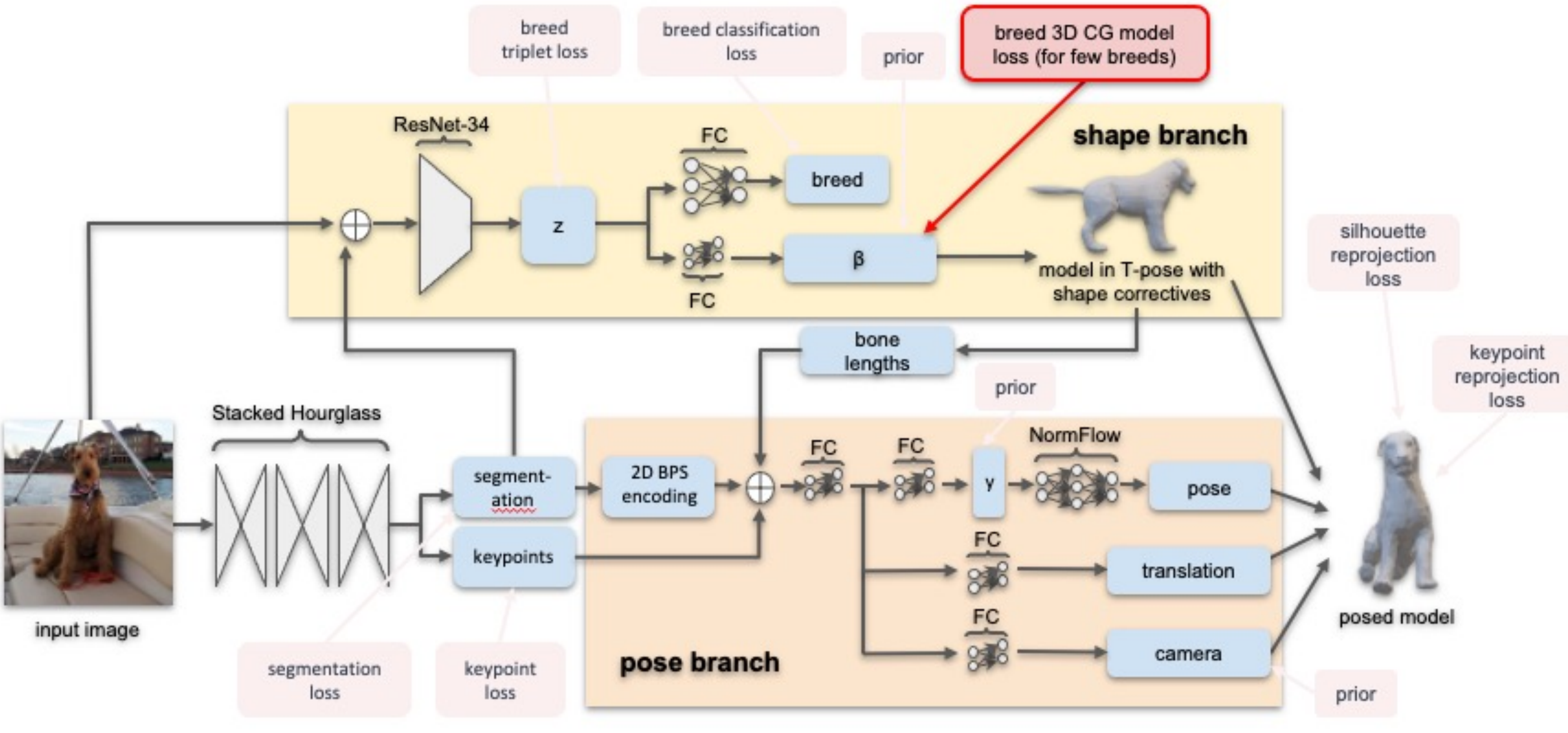
Breed-aware reconstruction



Breed-aware reconstruction



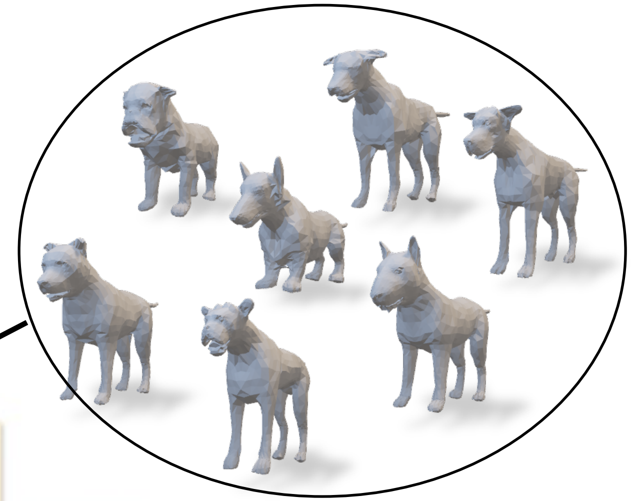
Breed-aware reconstruction

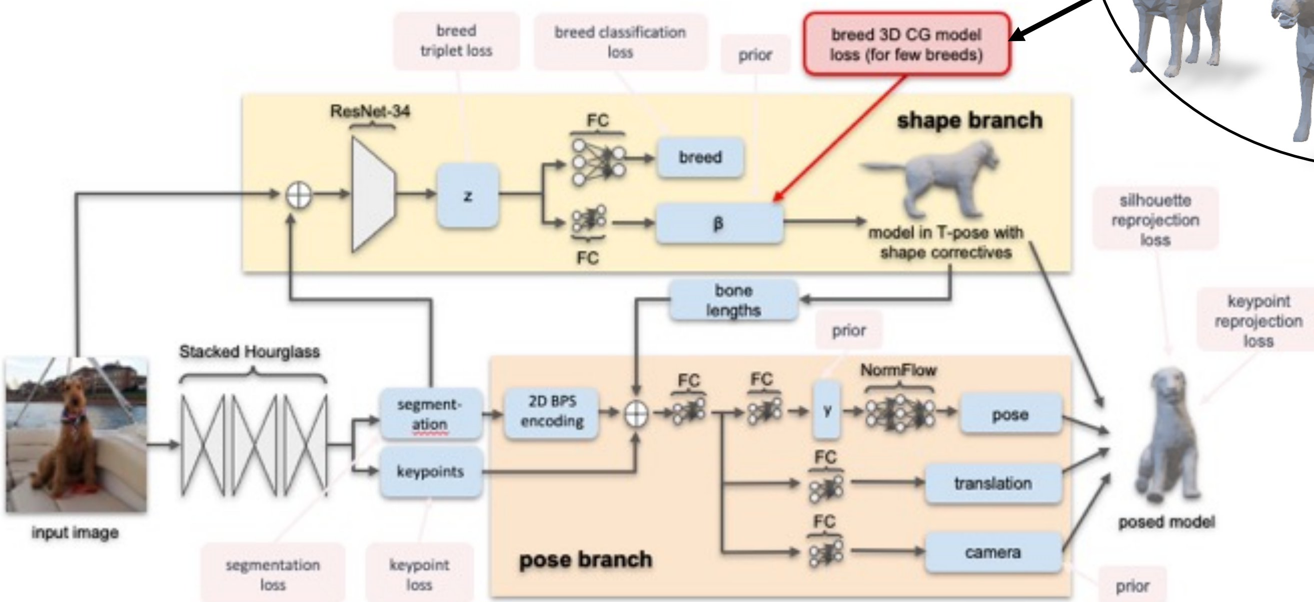


$$L_{3D}^{B} = (\beta_{pca}^{pred} - \beta_{pca}^{breed})^2 + (\kappa^{pred} - \kappa^{breed})^2$$

Breed-aware reconstruction

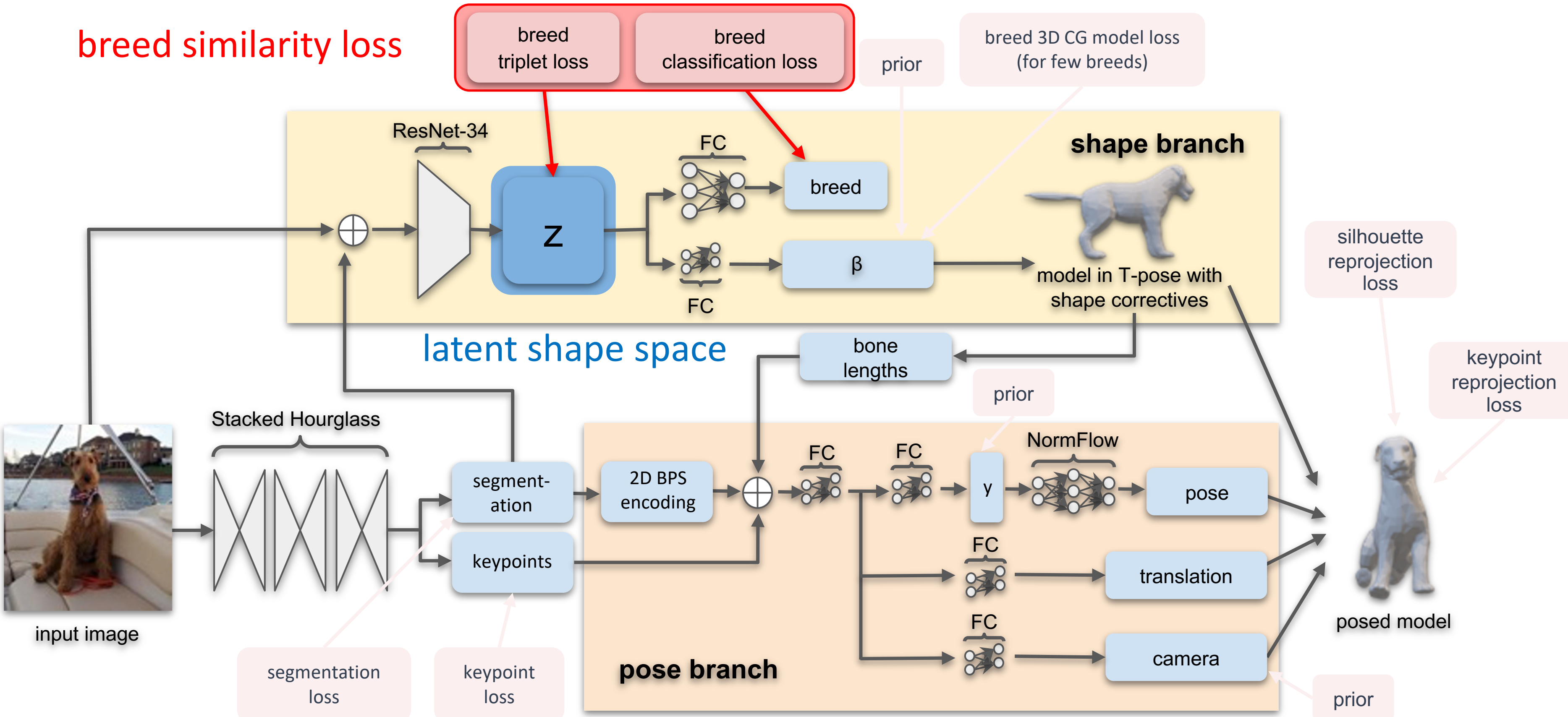
A small set of prototype 3D shapes to teach the network fine-grain breed details



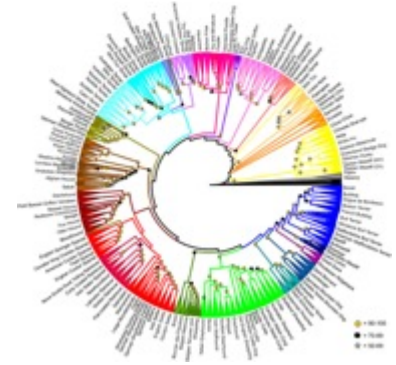


$$L^{\mathrm{B}}_{\mathrm{3D}} = (\beta^{\mathrm{pred}}_{\mathrm{pca}} - \beta^{\mathrm{breed}}_{\mathrm{pca}})^2 + (\kappa^{\mathrm{pred}} - \kappa^{\mathrm{breed}})^2$$

Breed-aware reconstruction



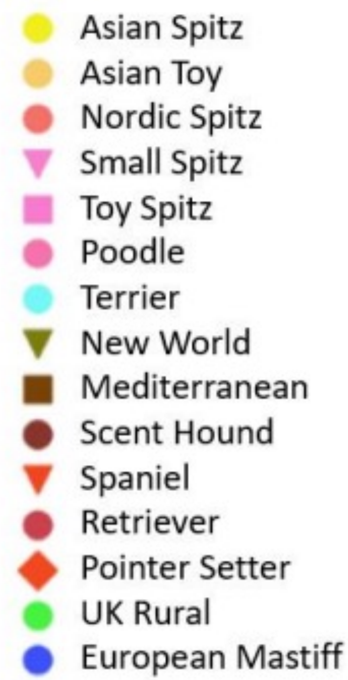
Latent space



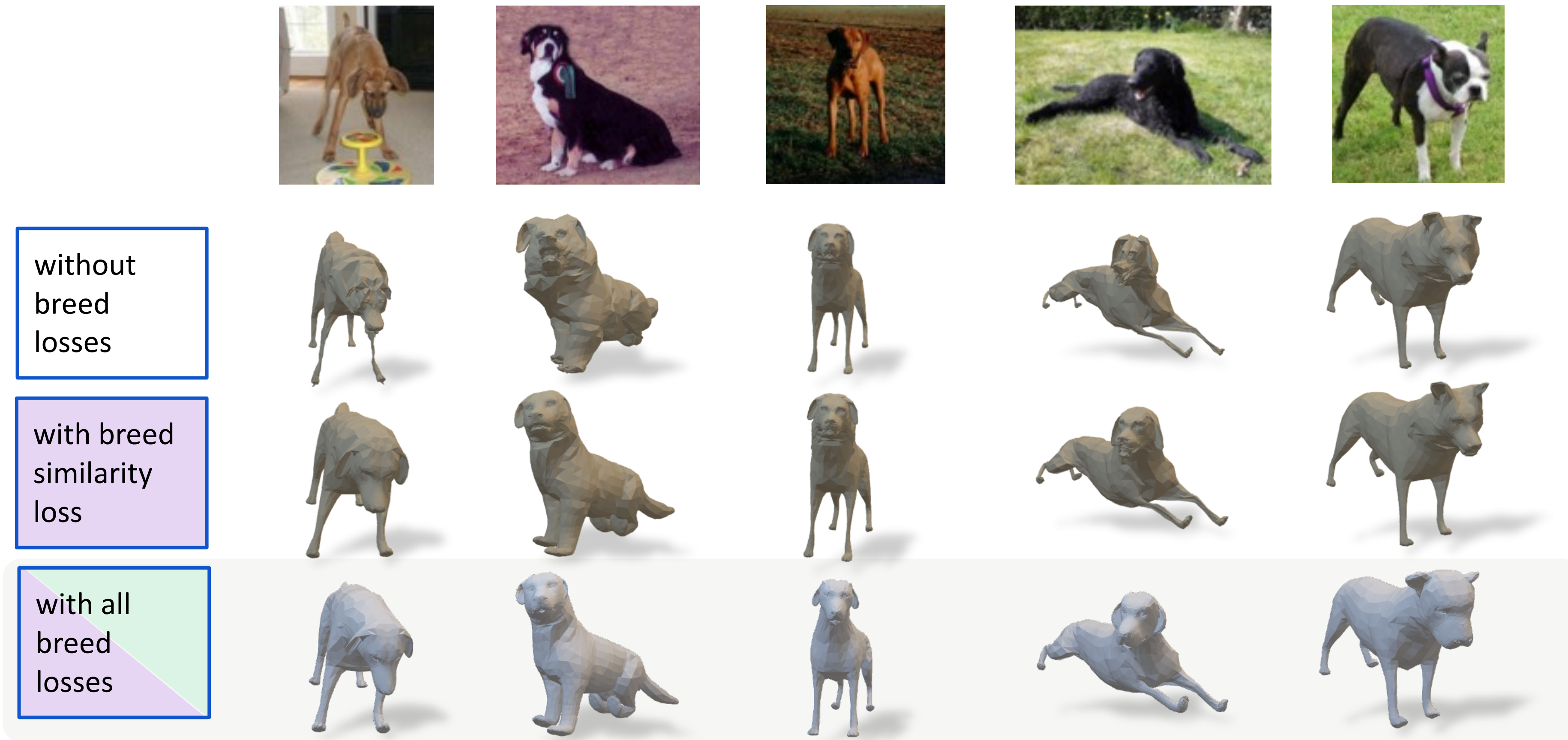
T-SNE plots of the latent variable z

without breed similarity loss

with breed similarity loss



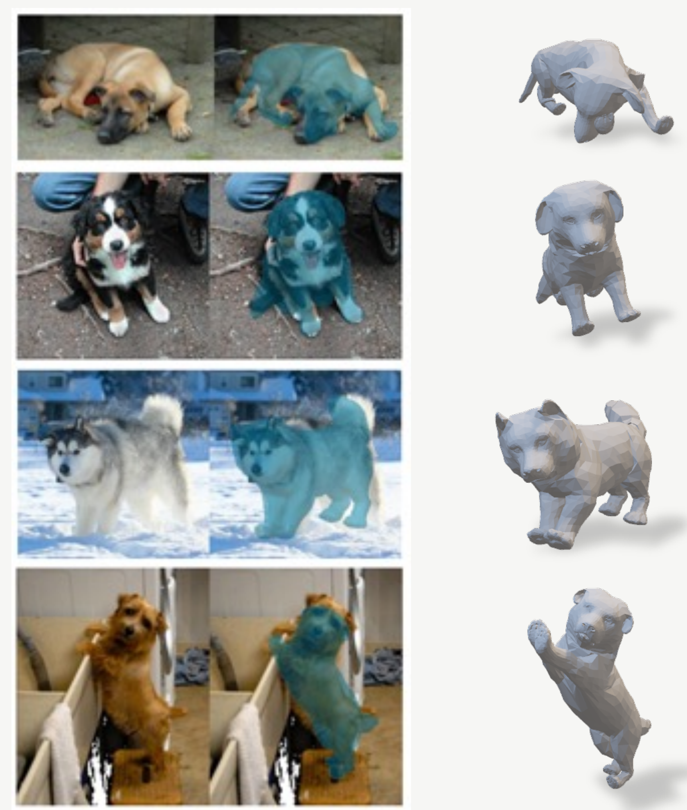
BARC results



BARC results

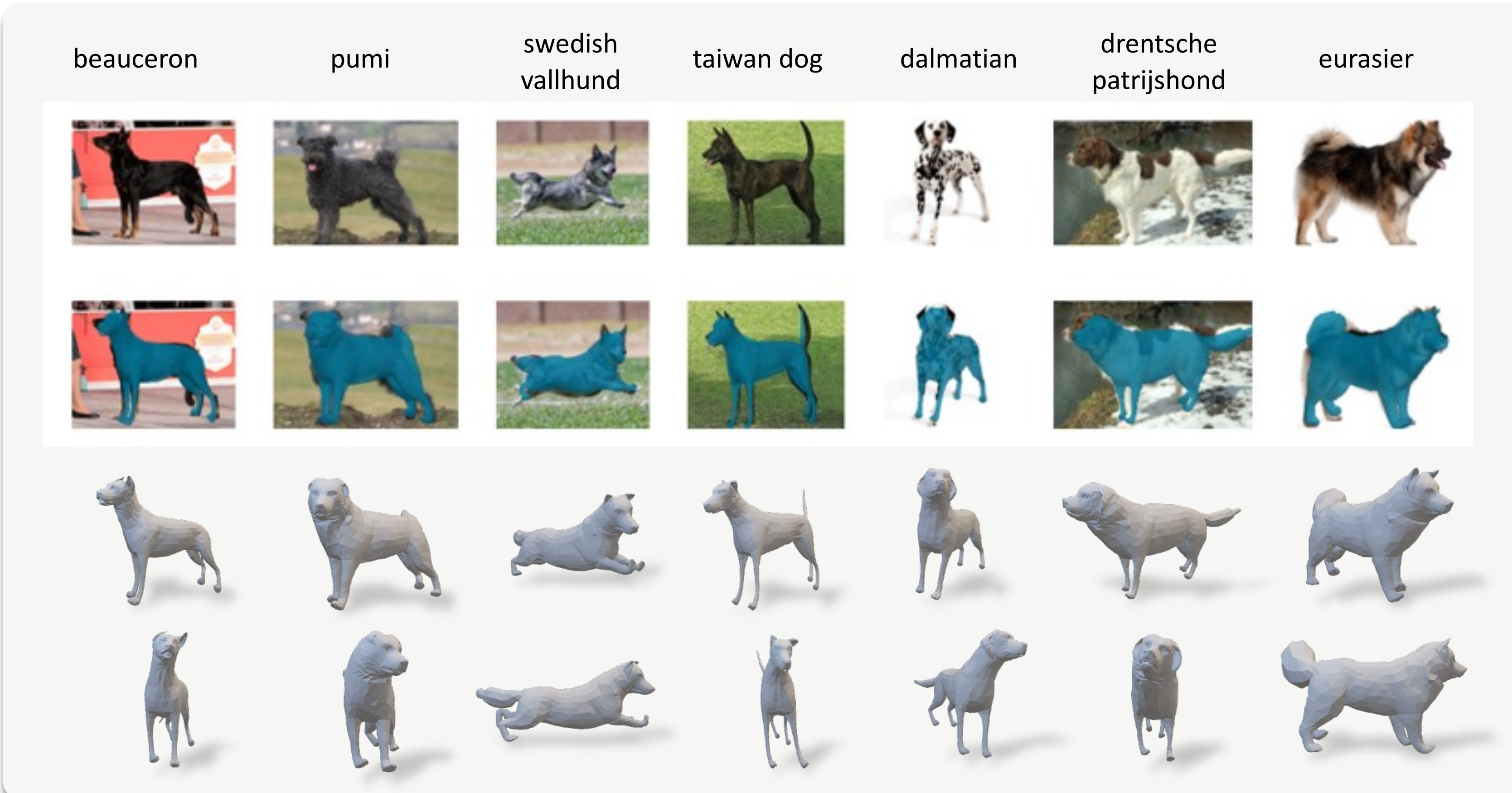
WLDO

BARC (ours)



BARC results

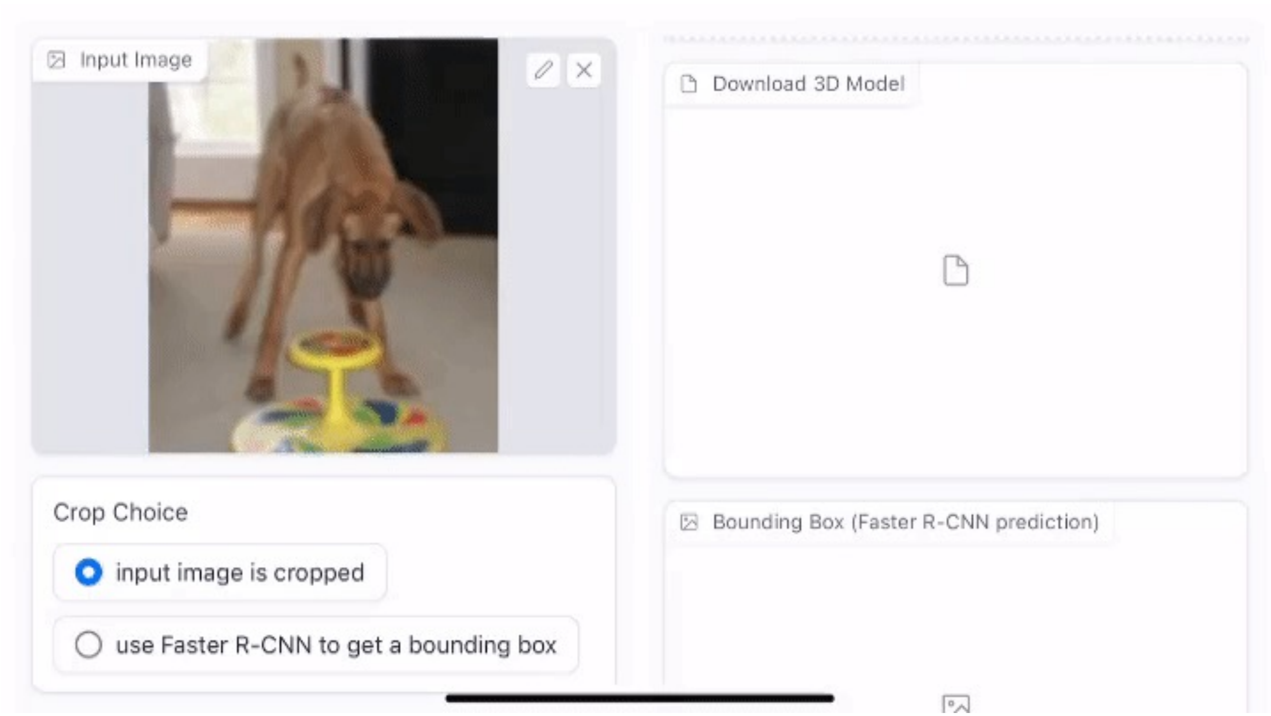
Unseen breeds



https://barc.is.tue.mpg.de/

Demo

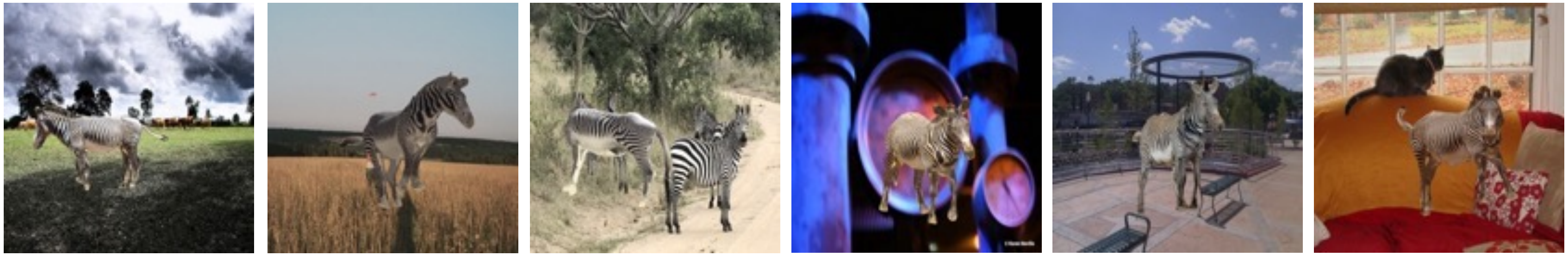
https://huggingface.co/spaces/runa91/barc_gradio



SMALST

Predict 3D shape, pose and texture of the Grevy's zebra from images

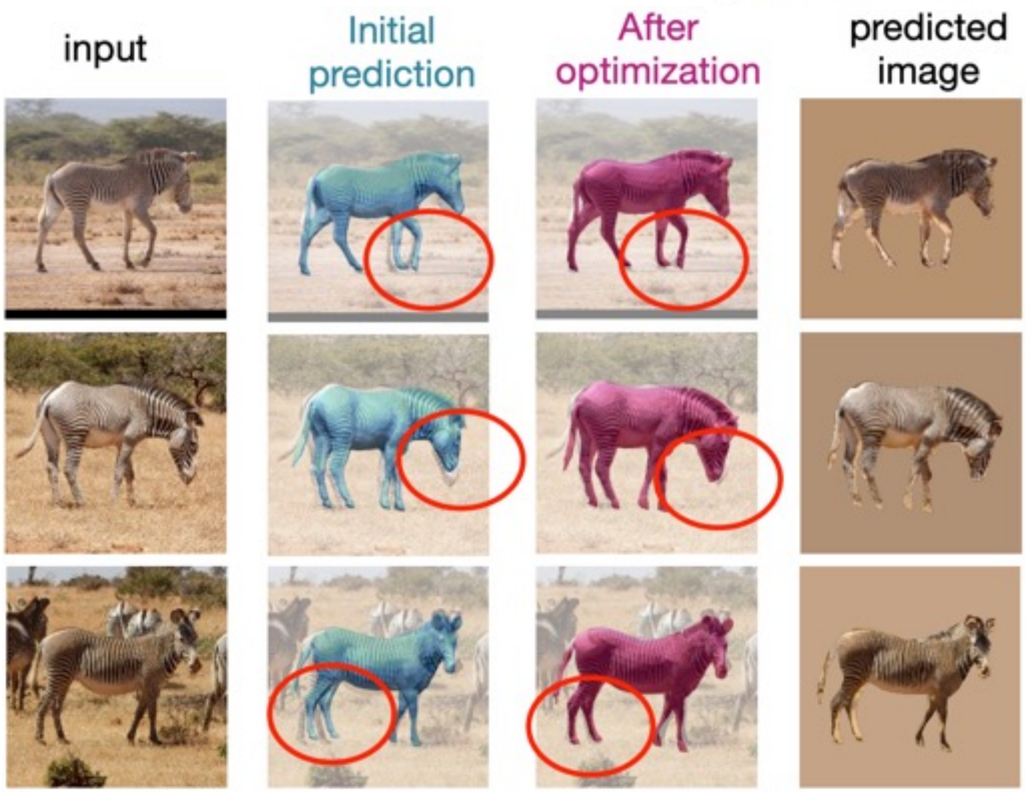
Trained on a synthetic dataset created with a set of SMAL zebra avatars

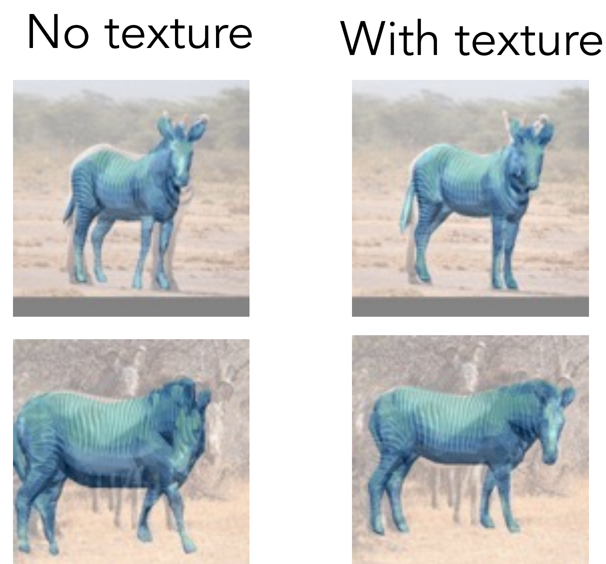


S. Zuffi et al., 3D Safari: learning to estimate zebra pose, shape and texture from images in-the-wild, ICCV 2019

Results

Method	PCK@0.05	PCK@0.1
(A) SMAL (gt kp and seg)	92.2	99.4
(B) feed-forward on synthetic	80.4	97.1
(C) opt features	**62.3**	**81.6**
(D) opt variables	59.2	80.6
(E) opt features bg img	59.7	80.5
(F) feed-forward pred.	59.5	80.3
(G) no texture	52.3	76.2
(H) noise bbox	58.7	79.9





Thank you