DEEP FACE VERIFICATION FOR SPHERICAL IMAGES

Marcos Cirne[1], Fernanda Andaló[1], Rafael Dias[2], Thiago Resek[3],
Gabriel Bertocco[1], Ricardo Torres[1] and Anderson Rocha[1]

[1]Institute of Computing, University of Campinas (UNICAMP), Brazil
[2]Eldorado Institute, Brazil, [3]Motorola Mobility LLC, Brazil

UNICAMP

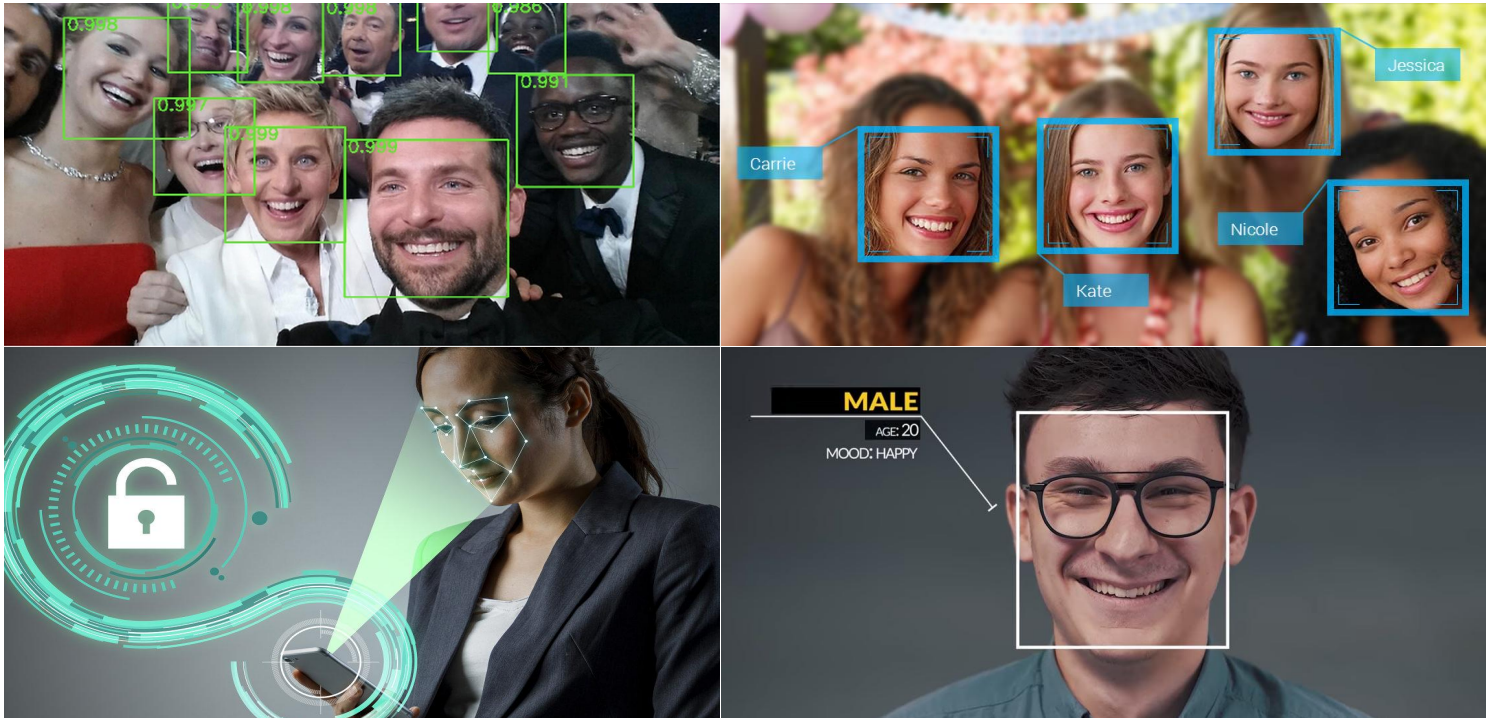
recod
reasoning for complex data

INTRODUCTION

UNICAMP

recod
reasoning for complex data

Several problems regarding face image analysis



Leveraged by the power of **CNNs**

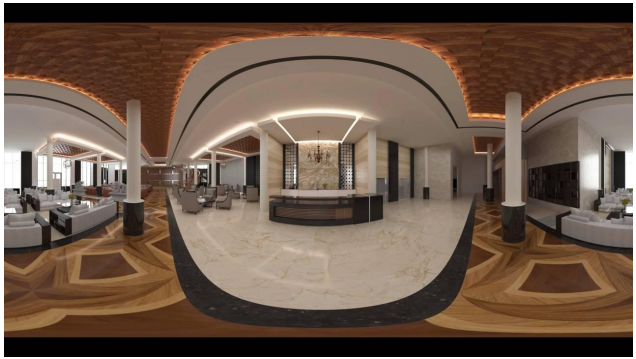
UNICAMP

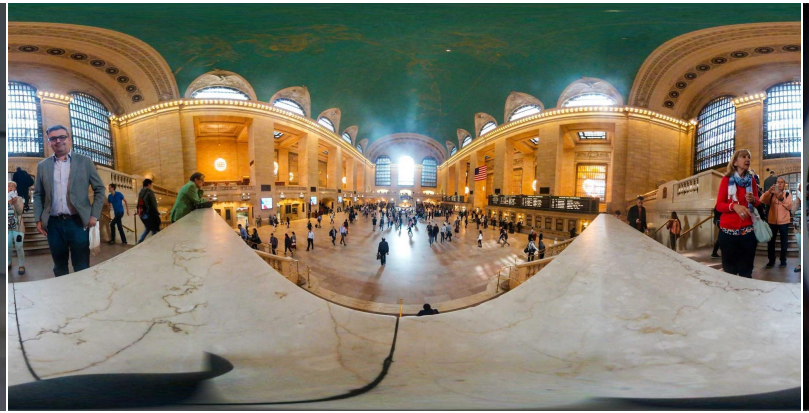
recod
reasoning for complex data

Popularity of 360° cameras is increasing

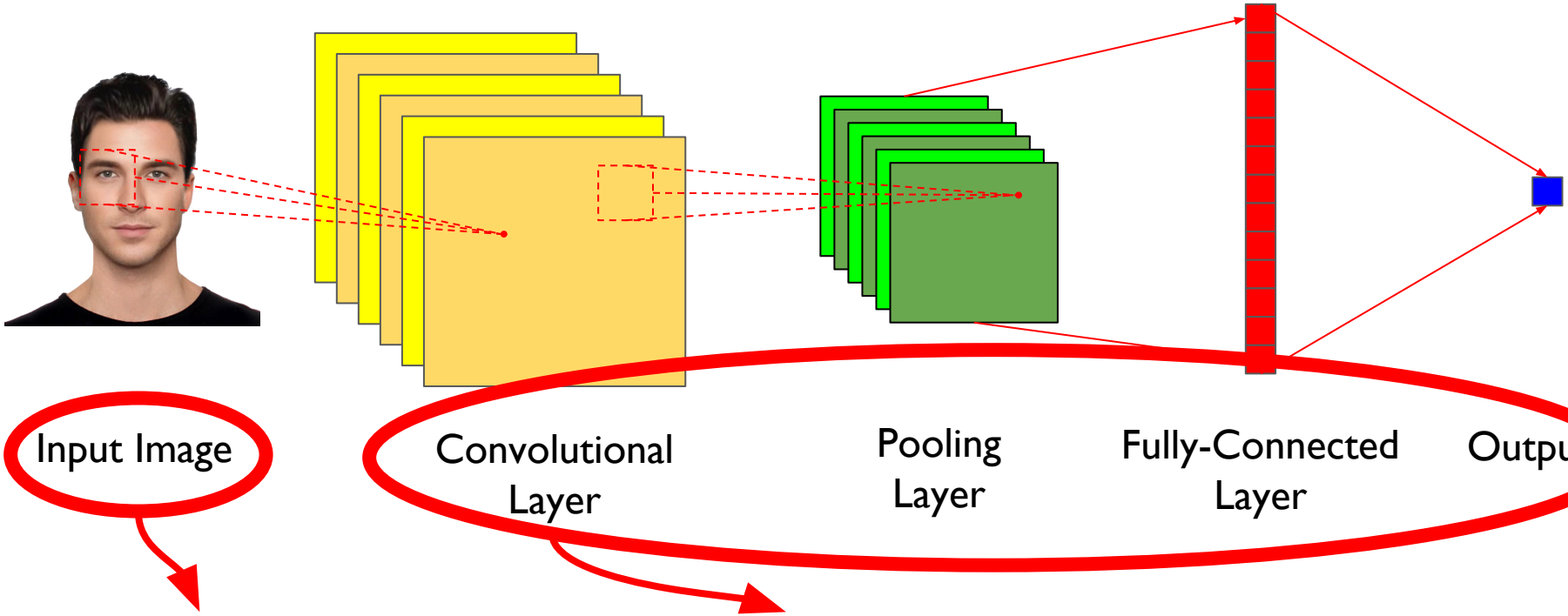
Wider field-of-view than traditional cameras

Examples of applications





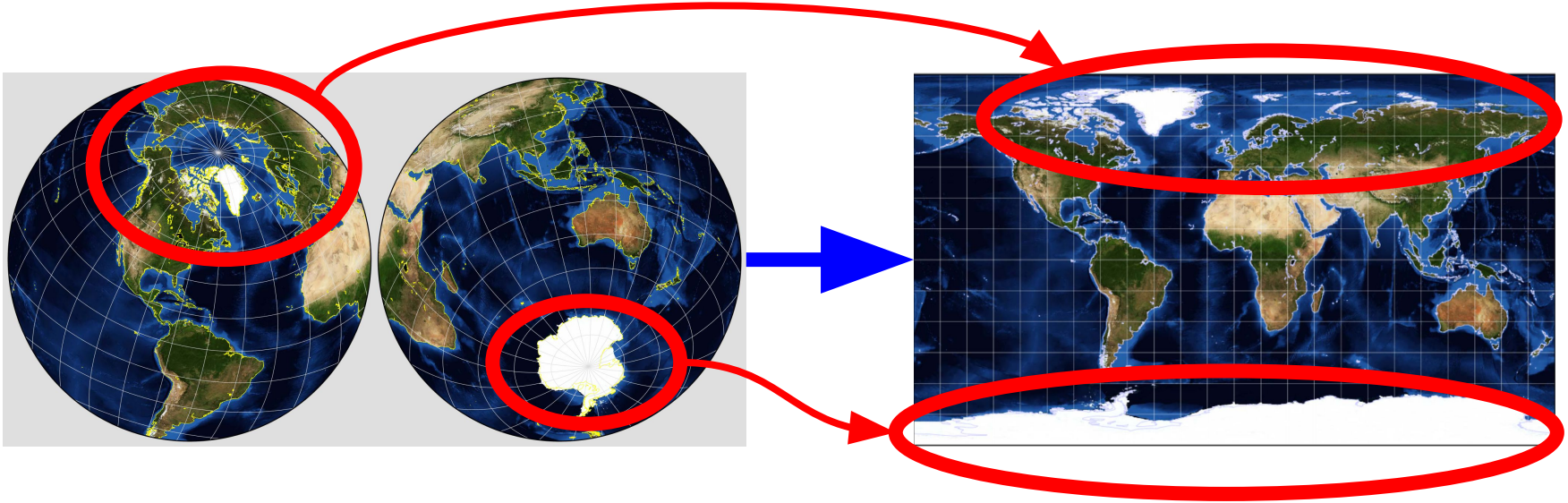
Problem: conventional CNNs are not trained for spherical images



UNICAMP

recod
reasoning for complex data

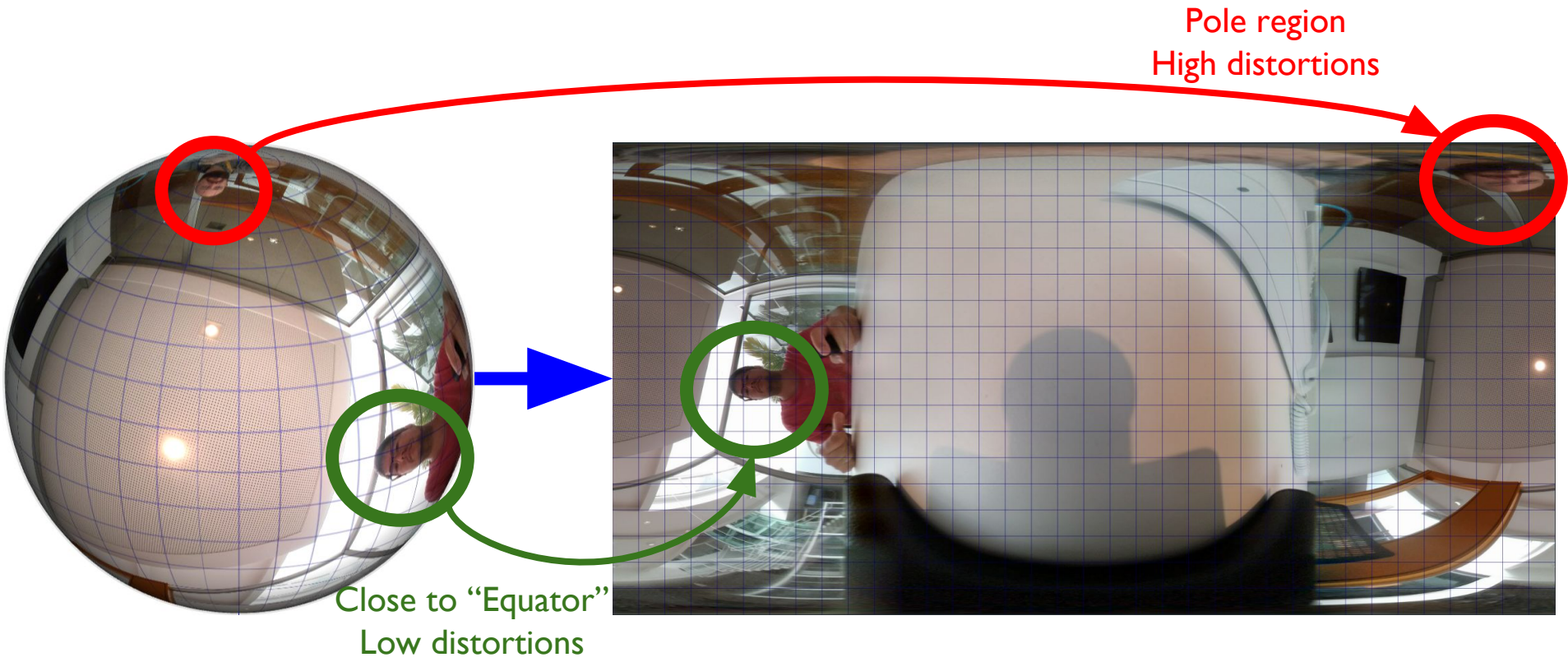
Solution: apply transformation from **polar coordinates** to **euclidean coordinates**



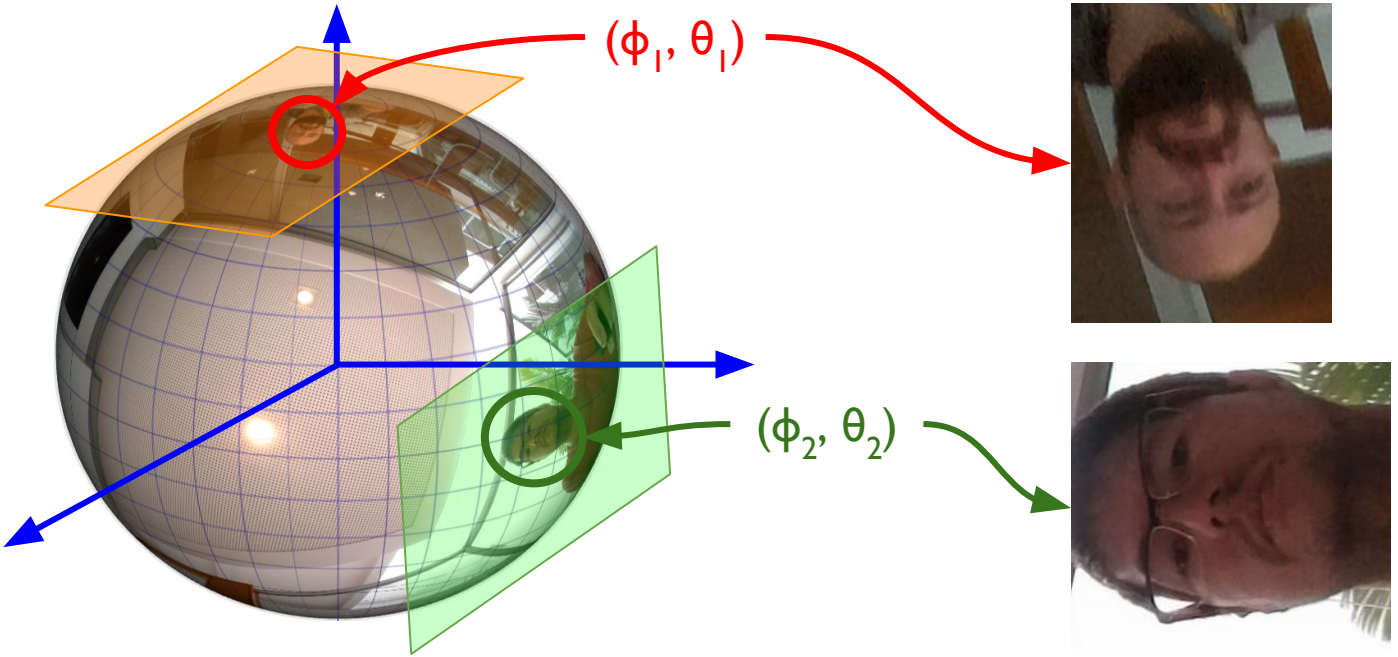
A possibility: **Equirectangular Projection**

Problem: Polar regions become severely distorted

The same happens with faces

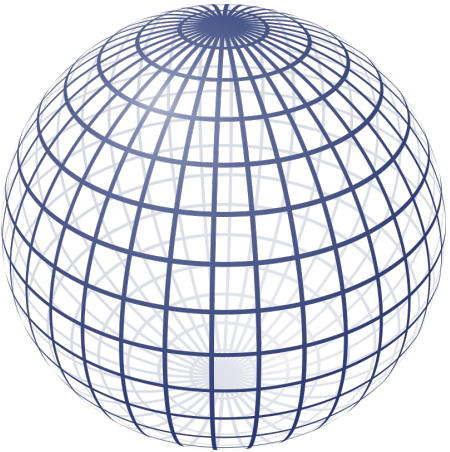


Naïve Solution: trace a tangent plane to every point (ϕ, θ) in a sampling process and calculate its planar projection



Computationally expensive!!!

CNNs can be trained to be immune to distortions



Spherical Image

sampling

Image 1 (ϕ_1, θ_1)	Image 2 (ϕ_2, θ_2)	Image 3 (ϕ_3, θ_3)
Image 4 (ϕ_4, θ_4)	•••	Image n (ϕ_n, θ_n)

Planar Projections

feedforward

CNN

But demands too many planar projections from a single spherical image

OBJECTIVES & CONTRIBUTIONS

OBJECTIVES & CONTRIBUTIONS

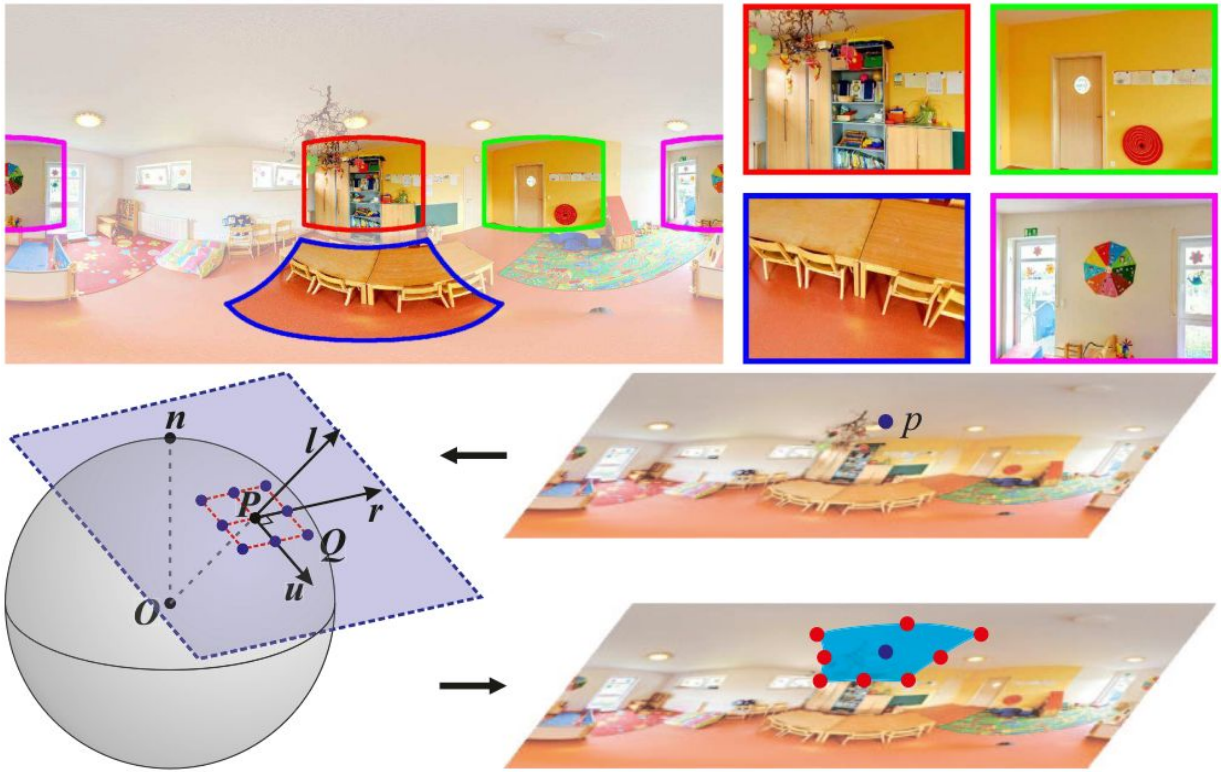
Novel approach for face verification which works on spherical images

New spherical face dataset: **MOT-360 Face**

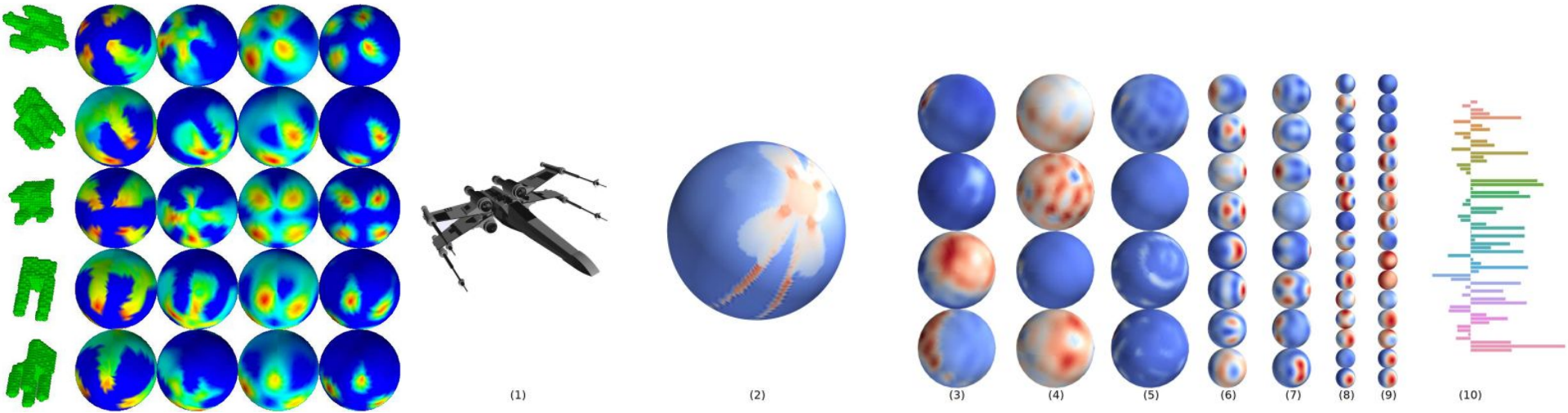
Spherical version of VGG Faces dataset: **VGG-360 Faces**

Comparative analysis against planar and spherical CNNs

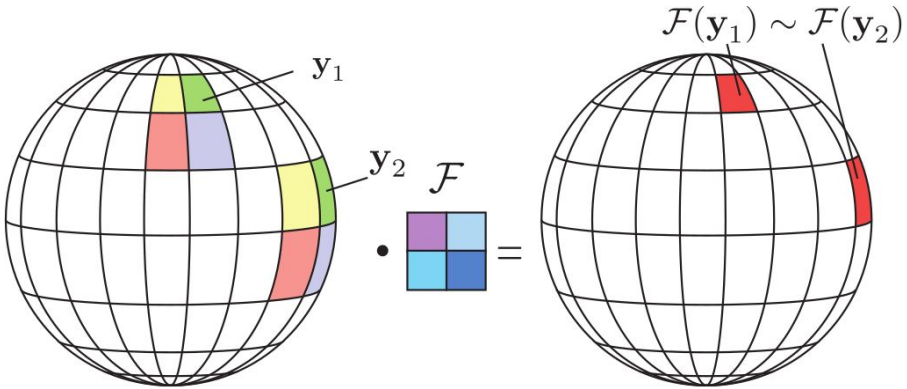
RELATED WORK

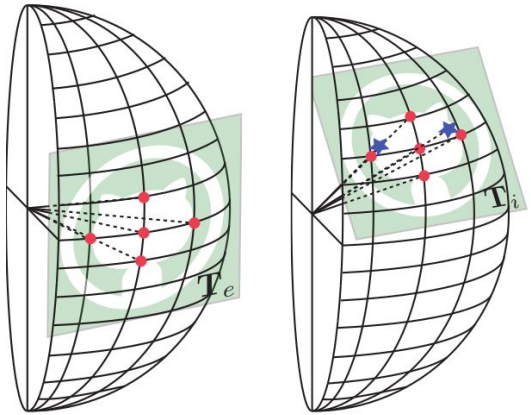


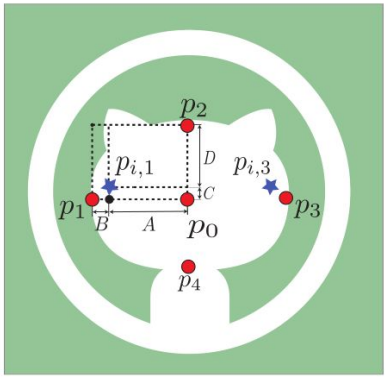
Zhao et al.: *Distortion-Aware CNNs for Spherical Images*; IJCAI 2018



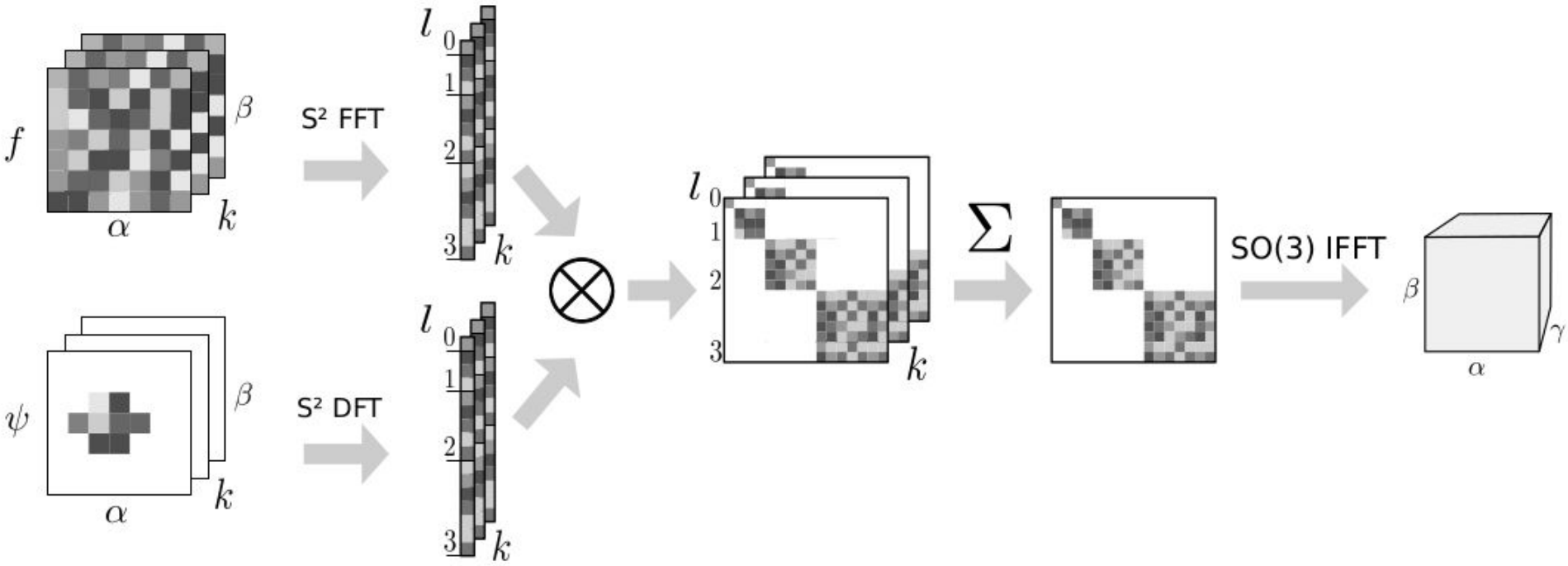
Esteves et al.: *Learning SO3 Equivariant Representations With Spherical CNNs*, ECCV 2018







Khasanova & Frossard: *Graph-Based Classification of Omnidirectional Images*, ICCV 2017



S2-CNN by Cohen et al.: *Spherical CNNs*, ICLR 2018

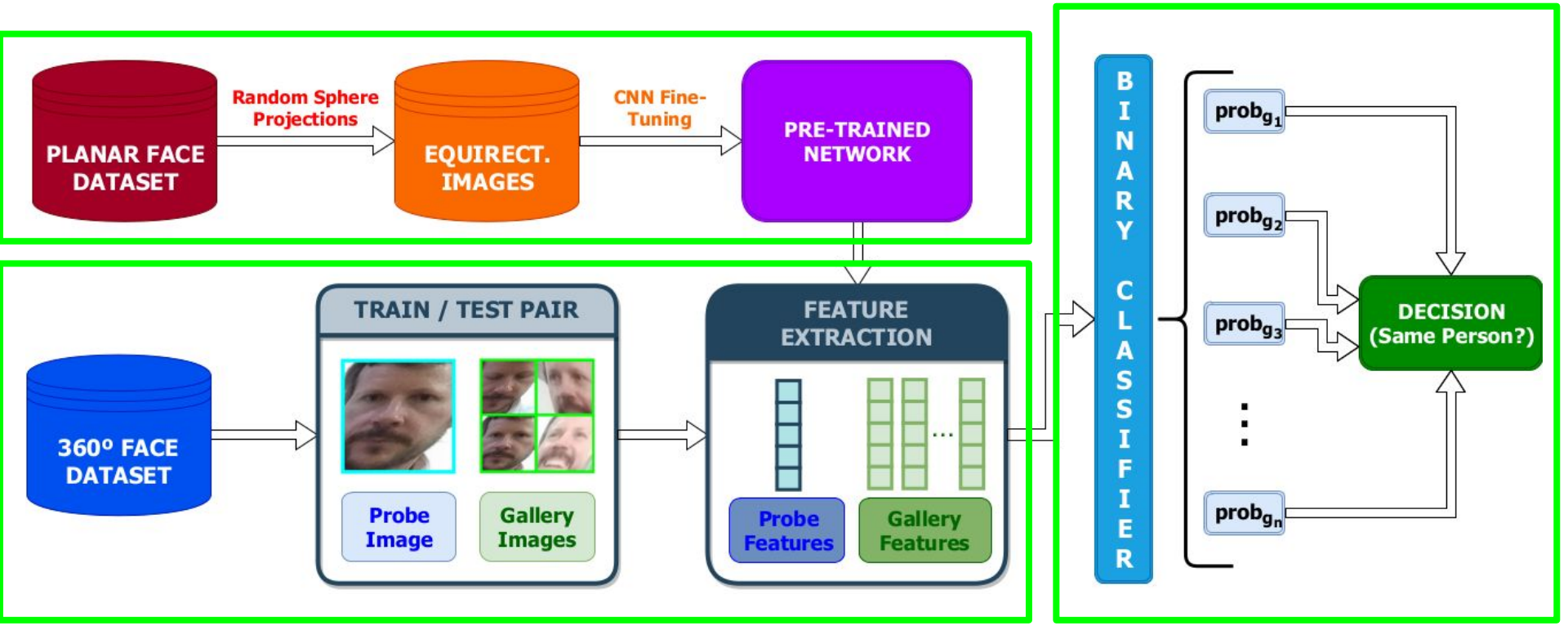
(a) 360° Cameras (b) 360° Image (c) Regular Kernel (d) SphereNet Kernel

Coors et al.: ***SphereNet***: *Learning Spherical Representations for Detection and Classification in Omnidirectional Images*, ECCV 2018

UNICAMP recod reasoning for complex data

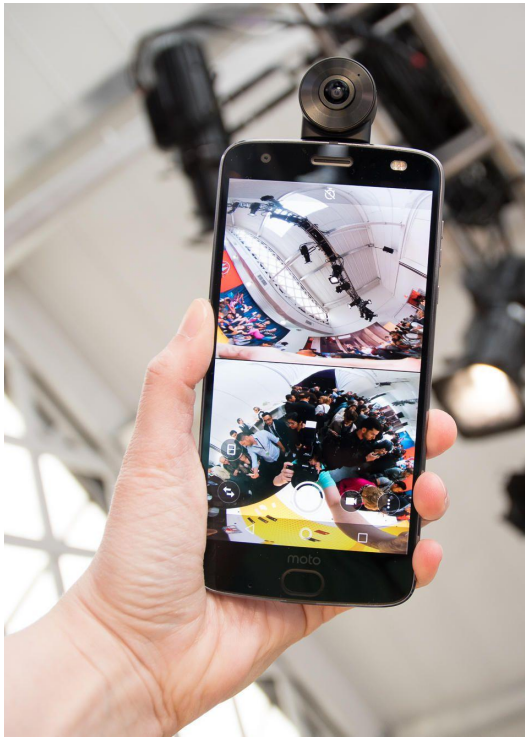
PROPOSED METHODOLOGY

Face verification methodology for spherical images

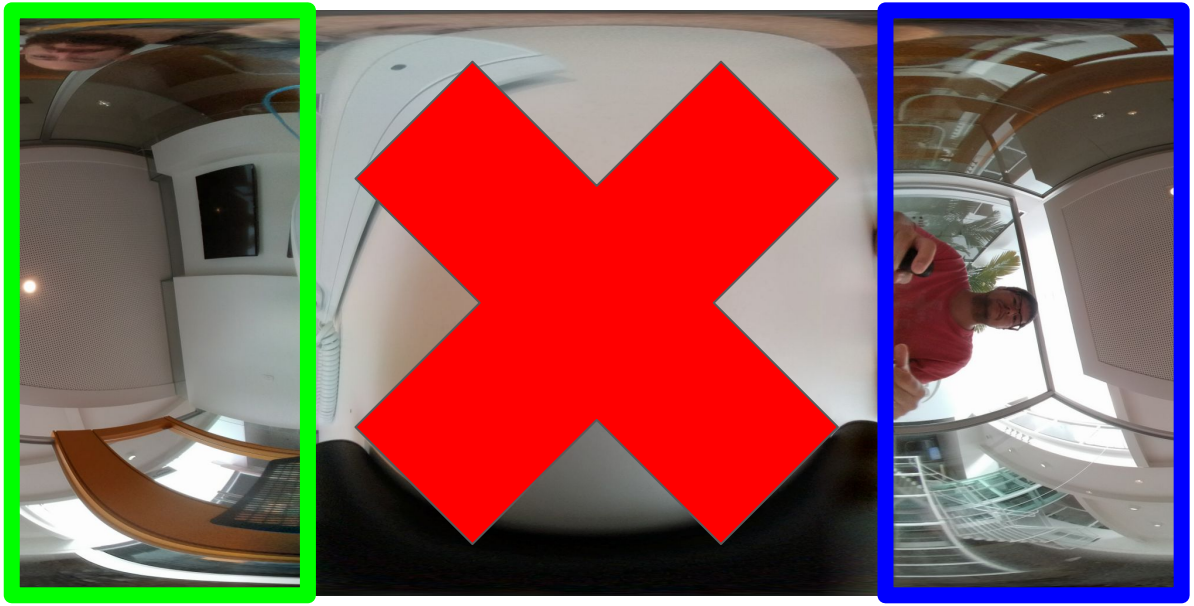


DATASETS

MOT-360 Camera Modulus



MOT-360 Dataset: Image Acquisition



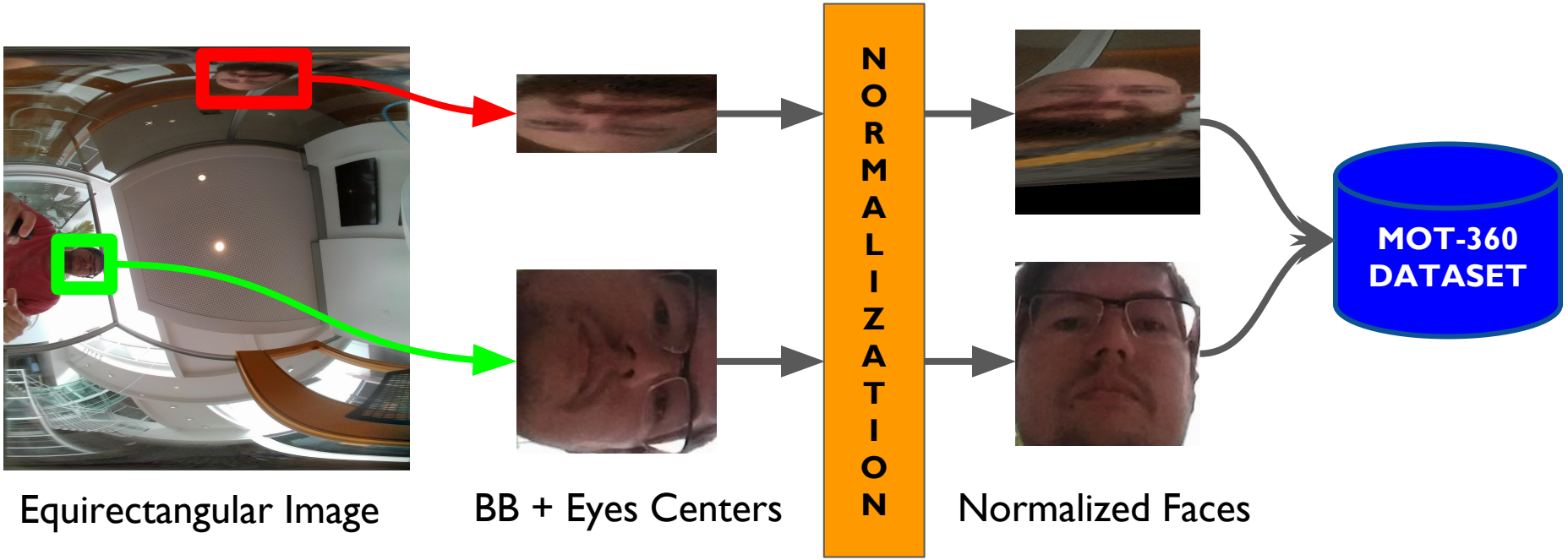
Original Equirectangular Image
(6240 x 3120 pixels)

Final Cropped Image
(3120 x 3120 pixels)

UNICAMP

recod
reasoning for complex data

MOT-360 Dataset: Face Annotation & Normalization



MOT-360 Dataset: General Information

360° Camera Angles: based on the polar coordinates of the bounding boxes of each annotated face

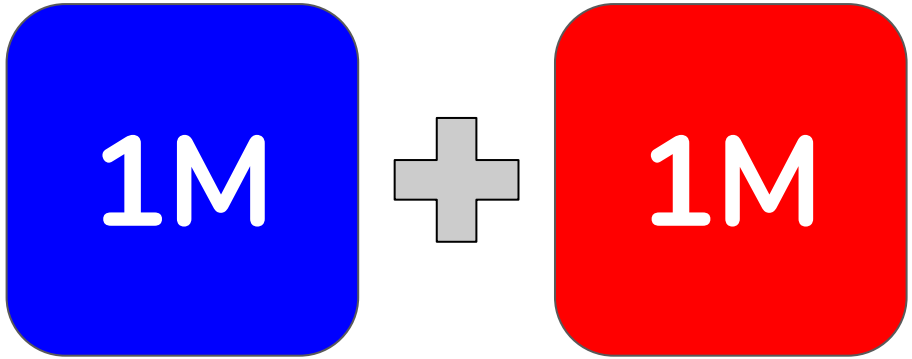
Dataset Size: **7,409** equirectangular face images from **52** unique individuals

Each individual has at least **25 face images**

MOT-360 Dataset: Training Protocol

Training Split	80% of the identities with the least number of images **42 IDs, 4128 images**
Test Split	All other images **10 IDs, 3281 images**

MOT-360 Dataset: Training Protocol



1 million positive + 1 million negative pairs created for each split

Gallery size: n = 10

MOT-360 Dataset: Examples

VGG 360 Face Dataset: General Information

Curated version of the VGG Face Dataset (2.6 M face images)

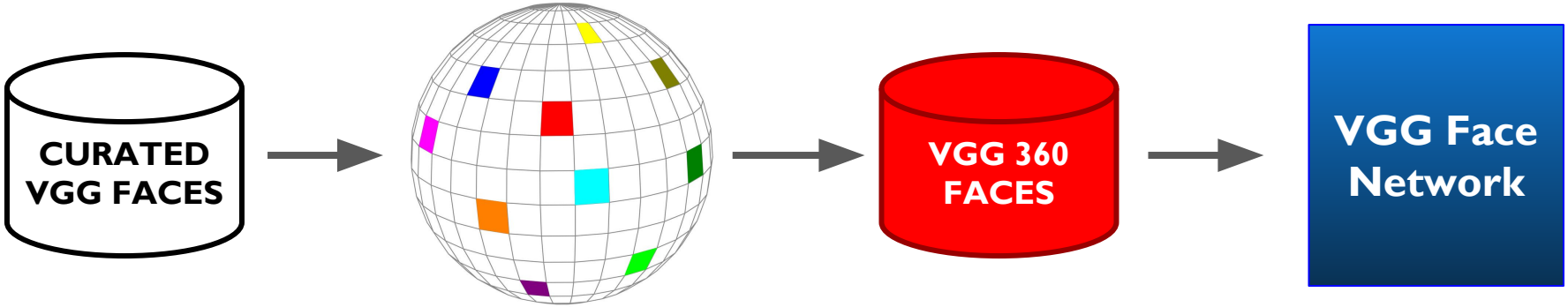
Dlib face detector was run to detect eye positions in all images (only worked for a small portion of the dataset)

Final subset size: **~750,000** images with **2,558** IDs

UNICAMP

recod
reasoning for complex data

VGG 360 Face Dataset: Augmentation Pipeline



Planar Face Images (750 K images)

Random Sphere Projections

10 projections per face (7.5 M images)

CNN Fine-Tuning (80-20 split)

VGG 360 Dataset: Examples

RESULTS

Tests with several networks as feature extractors:

VGG Face: trained with the original VGG Faces (planar images)

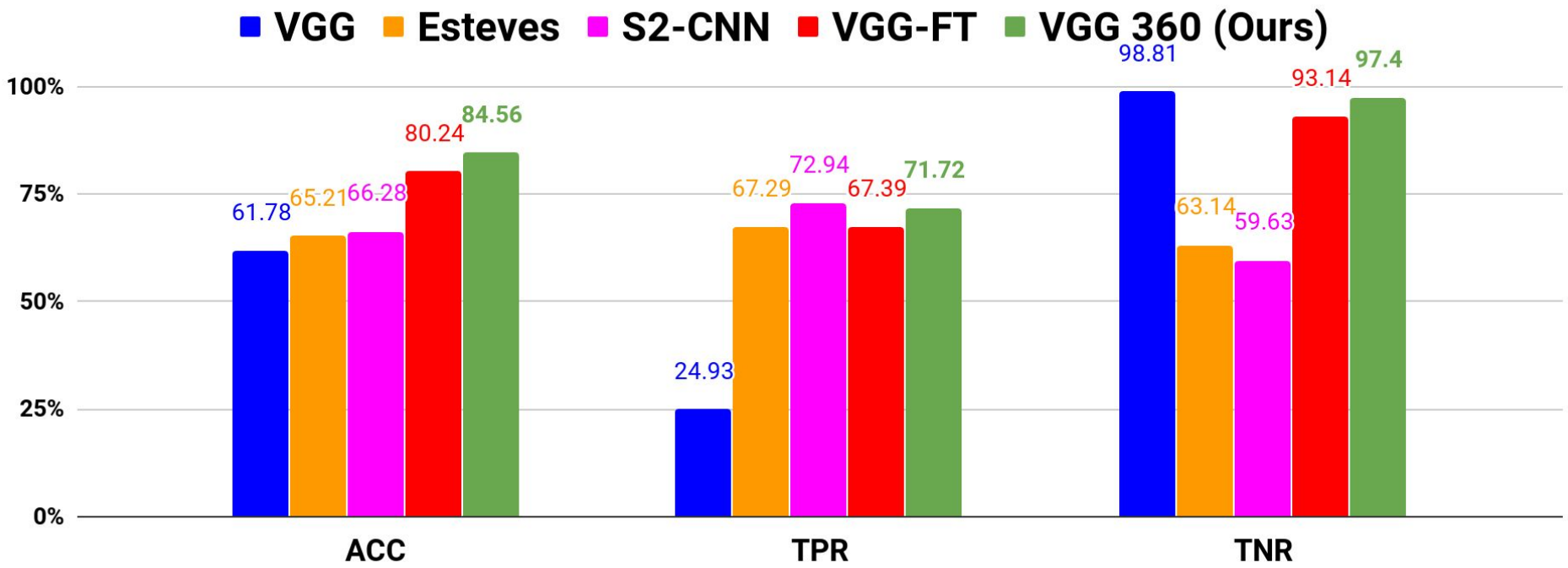
Spherical network proposed by **Esteves et al.**

S2-CNN: proposed by Cohen et al. and fine-tuned with VGG 360 Face dataset

VGG-FT: same as VGG Face, but fine-tuned with the curated version of VGG Face dataset (planar images)

VGG 360: our network, fine-tuned with VGG 360 Face dataset

RESULTS


VGG
Esteves
S2-CNN
VGG-FT
VGG 360 (Ours)
100%
75%
50%
25%
0%
61.78
65.21
66.28
80.24
84.56
24.93
67.29
72.94
67.39
71.72
98.81
63.14
59.63
93.14
97.4
ACC
TPR
TNR

CONCLUSIONS & FUTURE WORK

CONCLUSIONS

Feature Quality	Pair Protocol	Results	Distortion Issues
VGG-360 network can provide meaningful features for training a binary classifier for face verification	Significant number of positive / negative pairs for all the tests	Relevant results when compared against spherical CNNs	More research needed in operations on the spherical domain

FUTURE WORK

Extension	New Architecture	Comparative Analysis	Dataset Upgrades
Extension of the proposed method for face identification	Development of an original deep architecture for face verification	Comparisons with other spherical CNNs	Addition of more identities and images to the proposed dataset

ACKNOWLEDGEMENTS

UNICAMP

recod
reasoning for complex data

THANK YOU!!!

Contact: marcos.cirne@ic.unicamp.br

UNICAMP

recod
reasoning for complex data