



SPEECH GESTURE RECOGNITION

PROGRESS PPT 2

RIGVED SHIRVALKAR



CONTENT

01

COMPLETED

02

NEXT STEPS

03

TO DISCUSS

COMPLETED

01

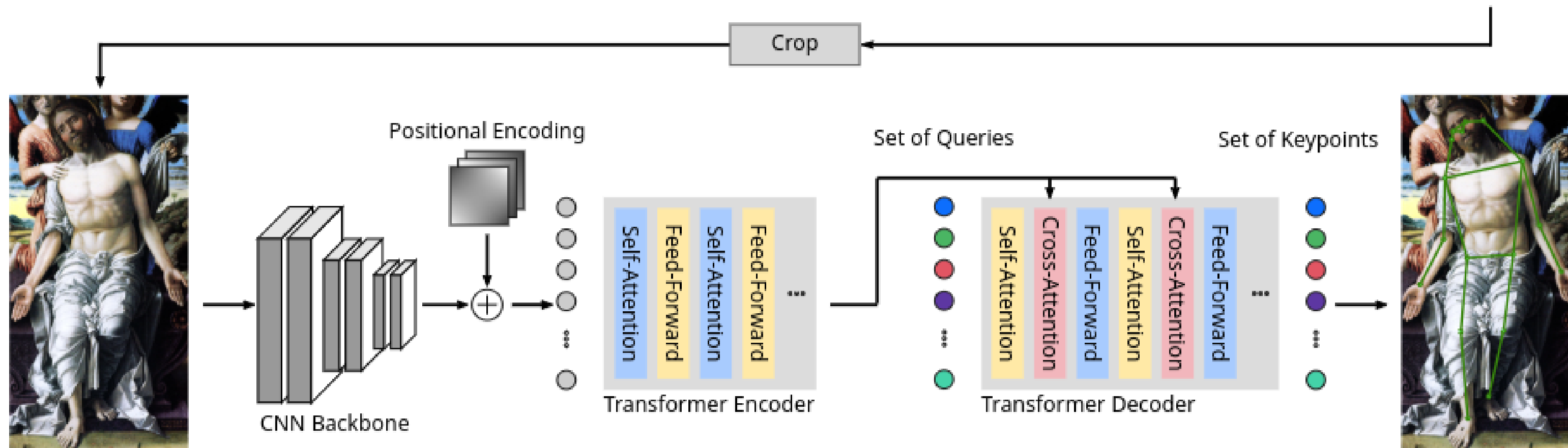
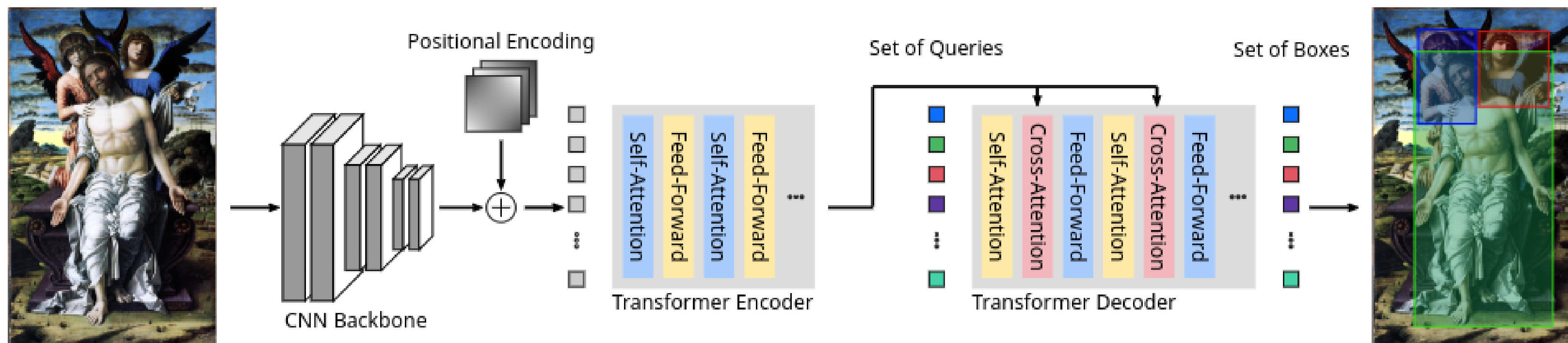
POSE DETECTION MODEL

02

POSE CLASSIFIER

POSE MODEL

- **MODEL USED:**
 - **POSE TRANSFORMER**
 - **[HTTPS://ARXIV.ORG/ABS/2207.02976](https://arxiv.org/abs/2207.02976)**
- **ADVANTAGES OF MODEL:**
 - **RECENTLY RELEASED AND OPEN SOURCE (GITHUB HAS NO LICENSE)**
 - **USES SEMI SUPERVISED LEARNING SO TRAINED ON LARGE DATA**
 - **TRAINED ON STYLE TRANSFER DATA AS WELL**
 - **CURRENT SOTA BY A SIGNIFICANT MARGIN**
 - **SEPARATES EACH PERSON FROM THE IMAGE AND DETECTS KEYPOINTS FOR EACH**



POSE CLASSIFIER

- **SIMPLE NEURAL NETWORK**
- **TRAINED ON KEY-POINT LABELS OBTAINED FROM POSE MODEL**
- **TBD: ALL POSSIBLE BLOCKS FOR THE BODY POSES**
- **CURRENT TRAINED BLOCKS:**
 - **BOTH HANDS RAISED**
 - **SINGLE HAND RAISED**
 - **NO HAND RAISED**
- **(WILL ADD HAND ON CHEST)**

Both Hands Raised



NEXT STEPS

- DECIDE SEPARATE BLOCKS FOR PALM GESTURES
 - PALM OPEN
 - PALM CLOSED
 - PALMS INTERLACED
 - POINTING (DIRECTION MATTERS TOO)
 - HOLDING THINGS
- CREATE DATA FOR PALM MODEL:
 - OBTAIN EACH PERSON'S IMAGES FROM PARTH'S DATA
 - THEN MANUALLY LABEL THE GESTURE BLOCKS
- TRAIN THE PALM DETECTION MODEL (YOLO V8)
 -

TO DISCUSS:

- FINALISING THE SCENE AND THE METHOD FOR TRAINING THE PALM DETECTION MODEL
 - MARY
 - PRETRAIN FOR PALM DETECTION OR DIRECT TRAIN
- THE CLASSES I CAN FINALISE FOR THE POSE CLASSIFIER
- THE CLASSES FOR THE PALM CLASSIFIER
- THE POSES I HAVE DECIDED FOR NOW:
 - BOTH HANDS RAISED (ORANS POSTURE)
 - HANDS CLASPED (PRAYER)
 - HANDS CROSSED ON CHEST (HUMILITY GESTURE)
- IDEA: WE CAN ALSO EXTRACT EACH PERSON FROM ALL IMAGES AND PERFORM UNSUPERVISED CLASSIFICATION OF BODY POSES