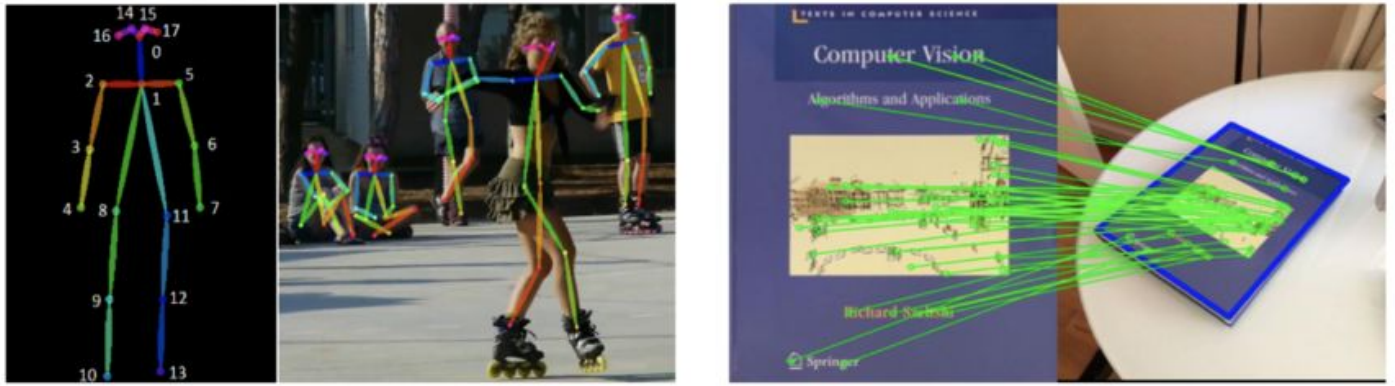
Human Keypoint Detection on Android

Roberto Scolaro - roberto.scolaro@studio.unibo.it

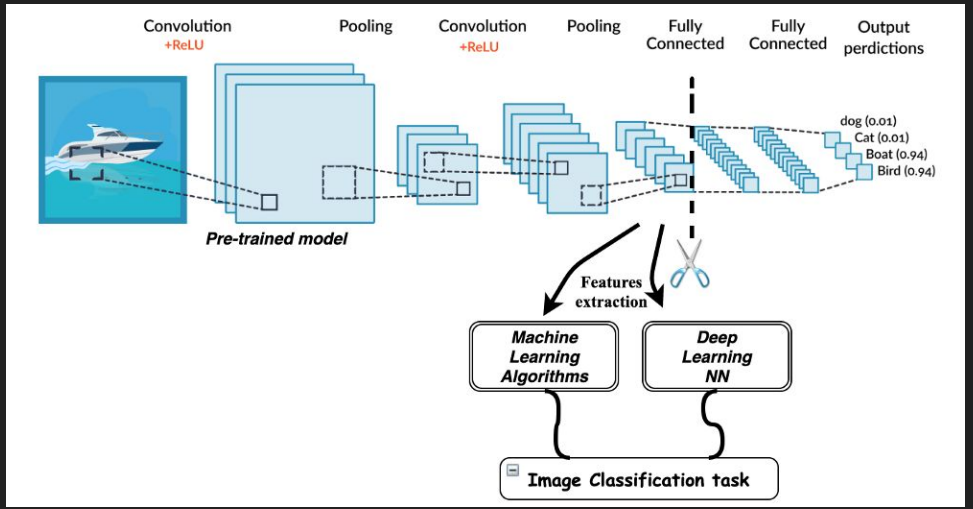
Analisi dell’obiettivo

Semantic Keypoints vs Interest Points



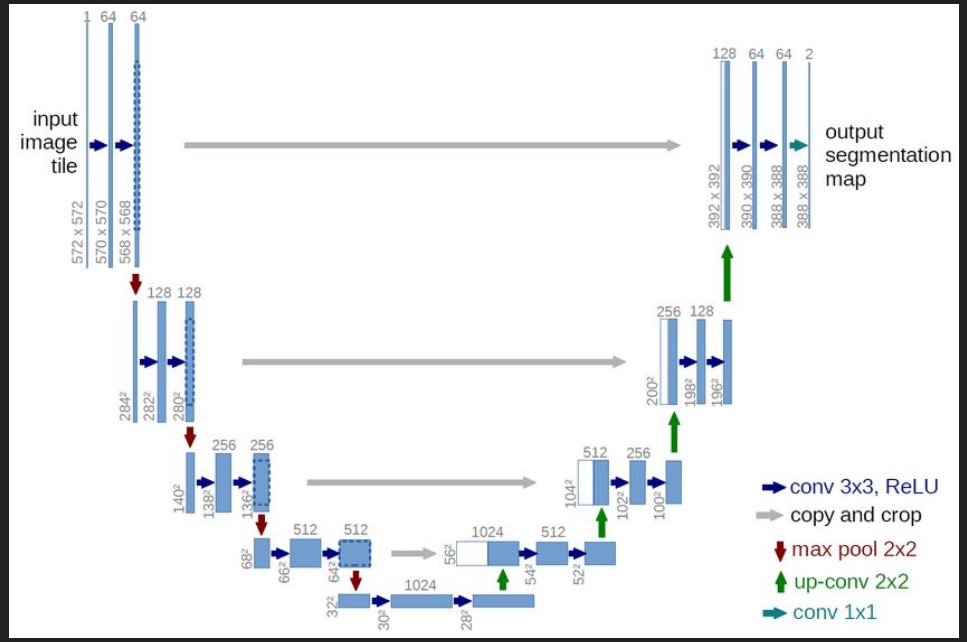
Transfer learning

- backbone con pesi di imagenet
- nuovi layer
- sfruttare feature già apprese per apprenderne altre
- ottenere direttamente le coordinate



U-Net

- fully convolutional neural network
- image segmentation
- 2 parti:
  - encoder
  - decoder



Heatmaps

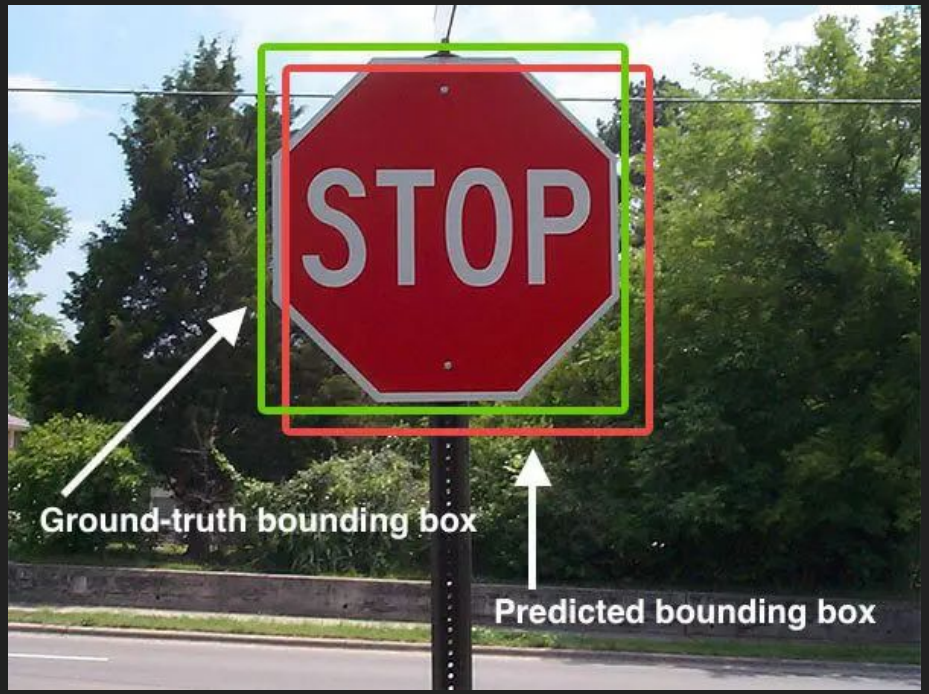
Creazione di una heatmap:

1. maschera di zeri delle dimensioni dell’immagine
2. keypoint = 1
3. GaussianBlur
4. normalizzazione

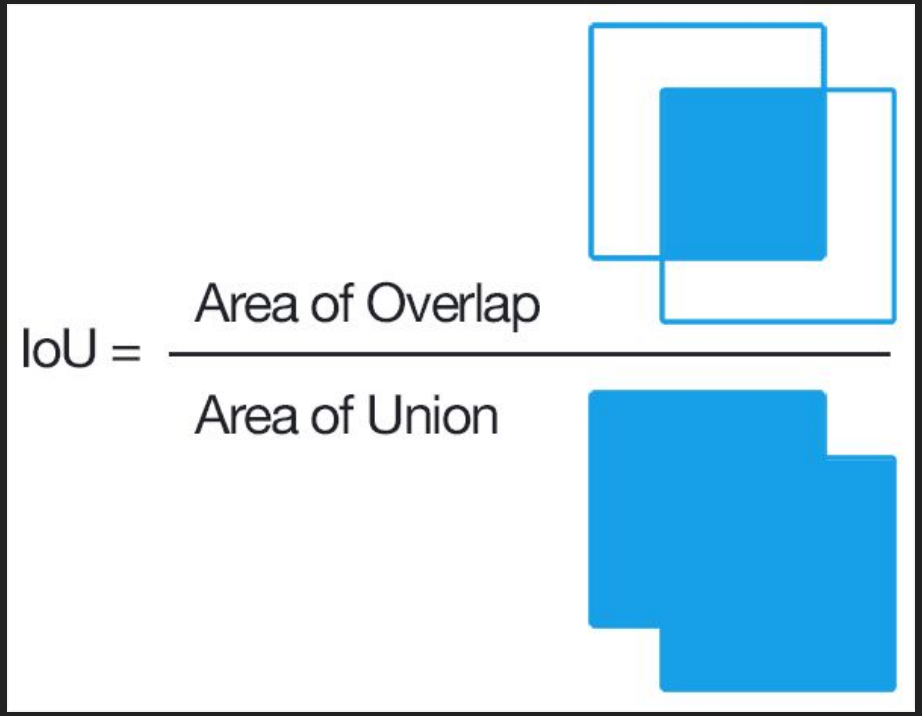
Loss function

IoU, Intersection over Union:

- image segmentation
- problema analogo

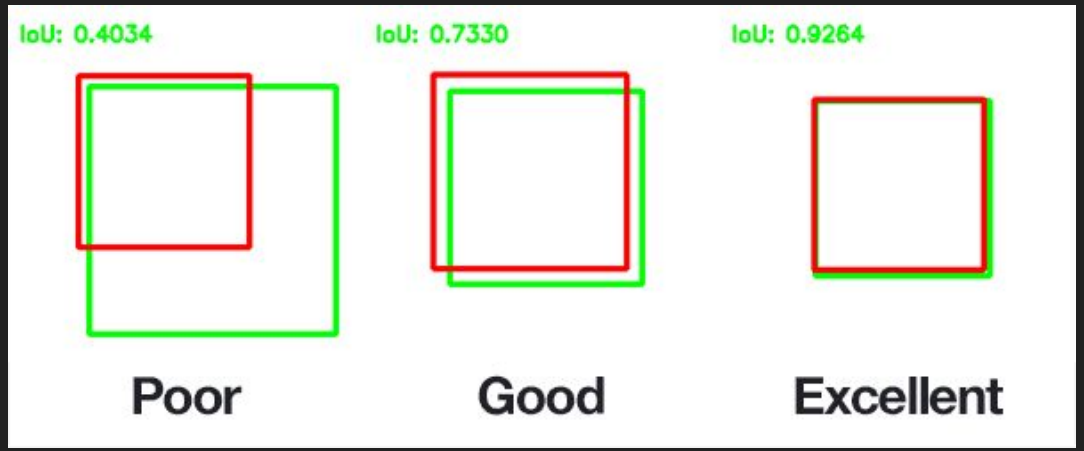


Loss function

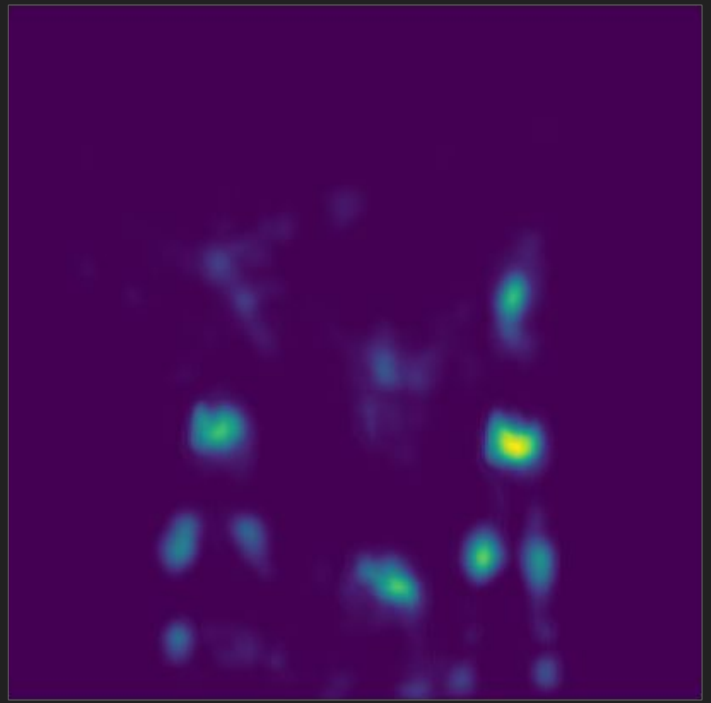


Loss function

3 casi possibili

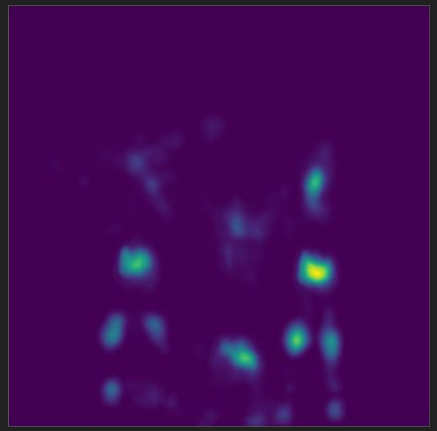


Risultati MPII

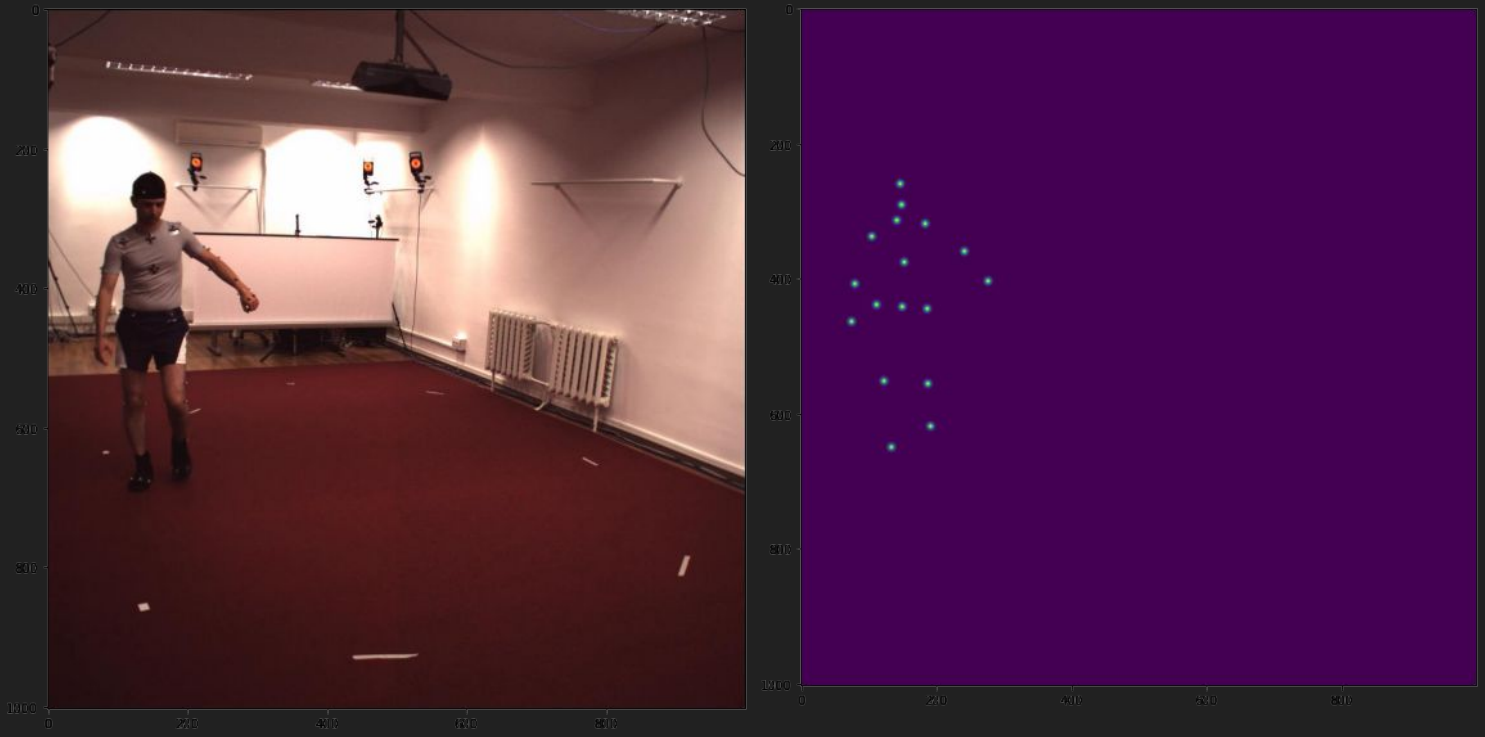


Risultati MPII

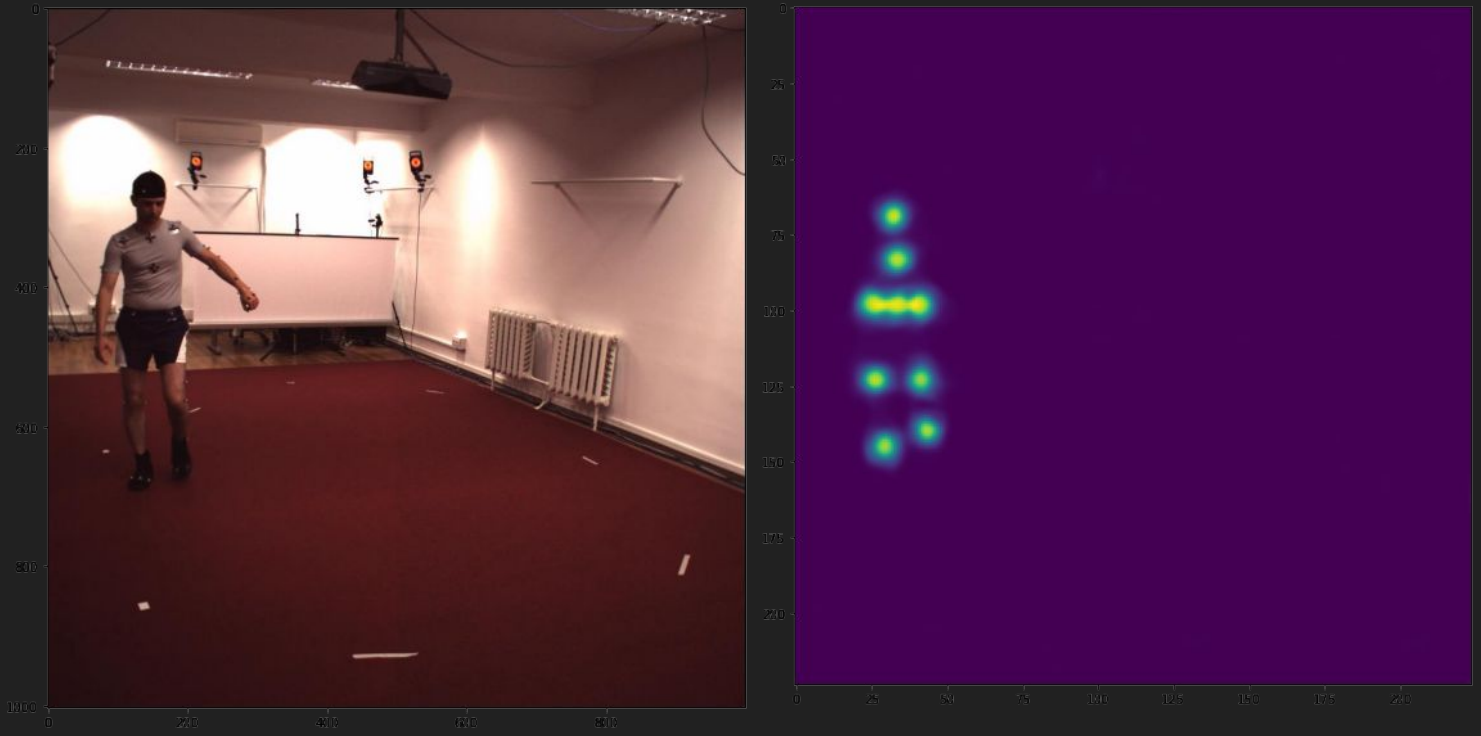
- allenato su single pose
- riconosce multipose
- prediction “in the wild” totalmente inaccurate



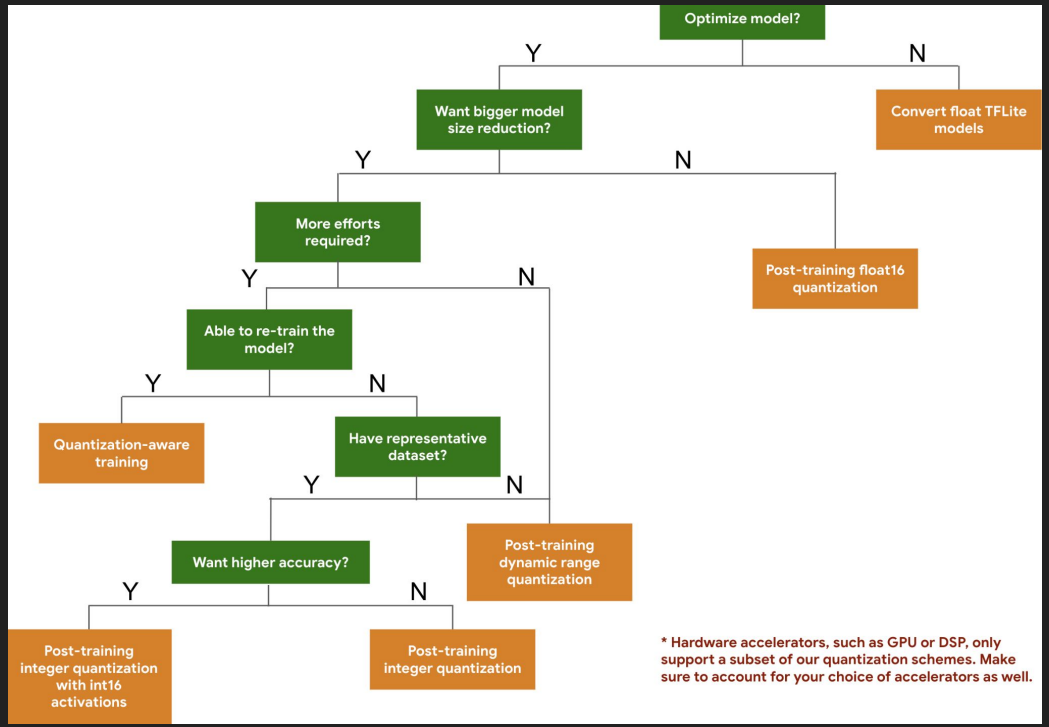
Risultati Human 3.6M



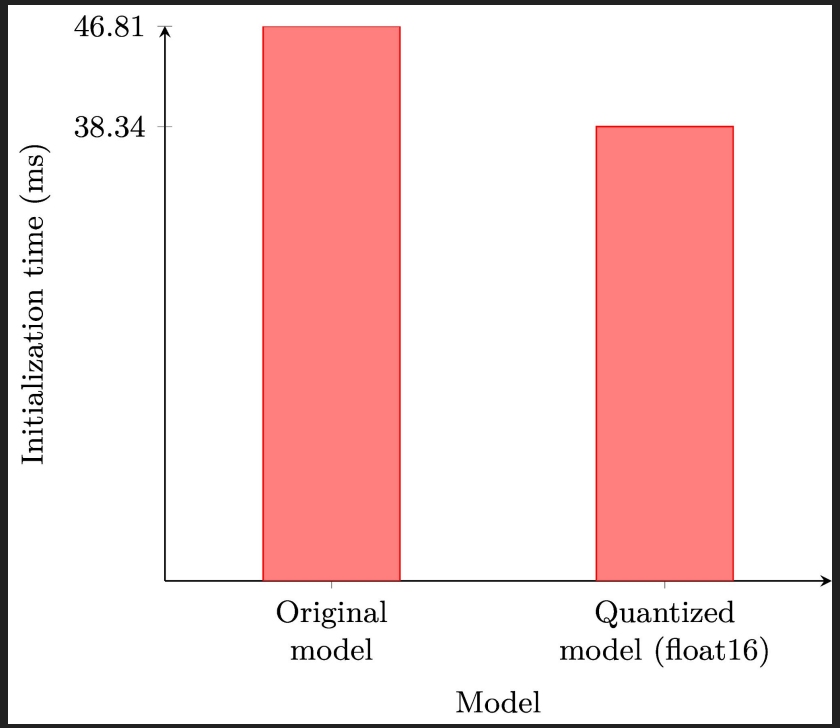
Risultati Human 3.6M

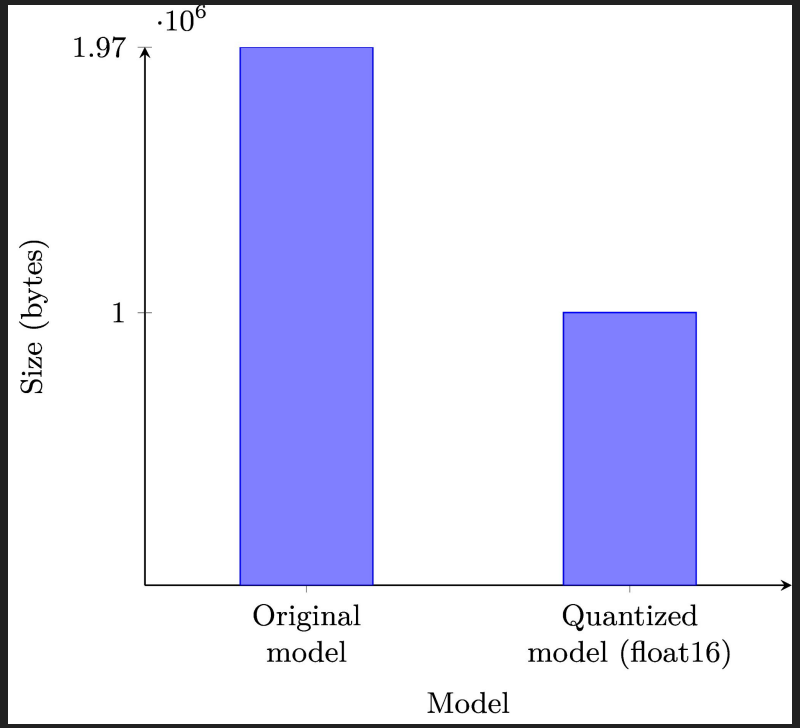


App Android

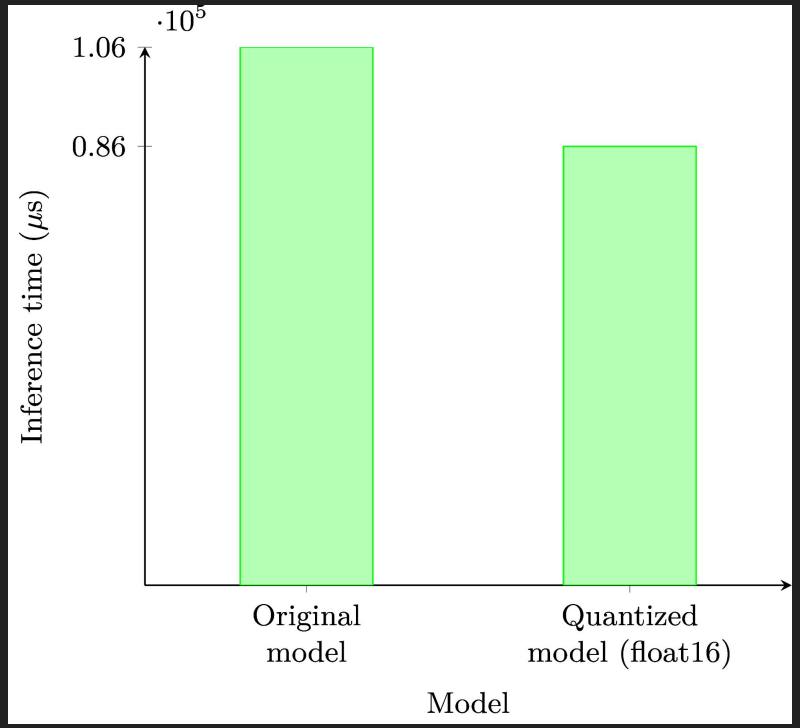

Optimize model?
Y
N
Want bigger model size reduction?
Convert float TFLite models
Y
N
More efforts required?
Post-training float16 quantization
Y
N
Able to re-train the model?
Y
N
Quantization-aware training
Have representative dataset?
Y
N
Want higher accuracy?
Post-training dynamic range quantization
Y
N
Post-training integer quantization with int16 activations
Post-training integer quantization
* Hardware accelerators, such as GPU or DSP, only support a subset of our quantization schemes. Make sure to account for your choice of accelerators as well.

App Android


46.81
38.34
Initialization time (ms)
Original model
Quantized model (float16)
Model


$\cdot 10^6$
1.97
1
Size (bytes)
Original model
Quantized model (float16)
Model

App Android


$\cdot 10^5$
1.06
0.86
Inference time (μs)
Original model
Quantized model (float16)
Model

Grazie per l’attenzione!