Detecting Sexually Provocative Images

Debashis Ganguly   **Mohammad H. Mofrad**   **Adriana Kovashka**
**University of Pittsburgh**

Motivation

❑ Overwhelming amount of visual data on the Internet
  ➤ E.g. each day 300 million photographs are uploaded to Flickr.
❑ Parents may want to restrict the visual contents their children can see.
  ➤ Automatic filtering of images for offensive contents is not perfect.
❑ Lots of manual effort is invested by digital content administrators to classify images in age restricted categories.
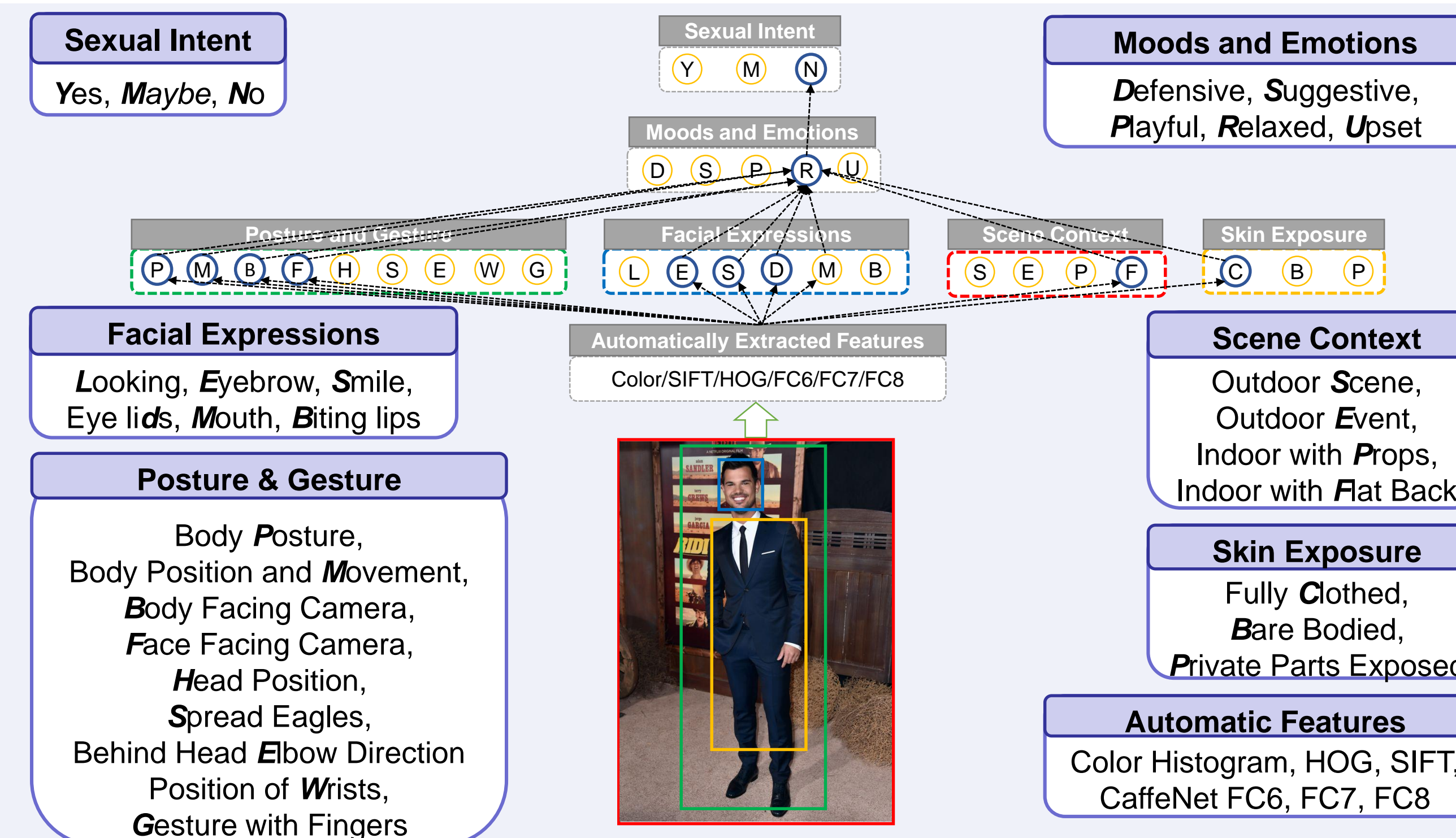
Limitations of Existing Approaches

❑ Existing approaches detect pornographic contents based on percentage of skin area exposed by the subjects in such images.
  ➤ Jiao et. al., "Detecting adult image using multiple features", Info-tech and Info-net 2001.
  ➤ Duan et. al., "Adult image detection method based on skin color model and support vector machine", Asian Conference on Computer Vision 2002.
  ➤ Zheng et. al., "Shape based adult image detection", International Journal on Image and Graphics 2006.
  ➤ Lee et. al., "Naked image detection based on adaptive and extensible skin color model", Pattern Recognition 2007.
❑ These works pay little attention to the intention behind the image composition and the goals of the photographic subject with respect to how the photo should be perceived.

(a)   (b)   (c)   (d)
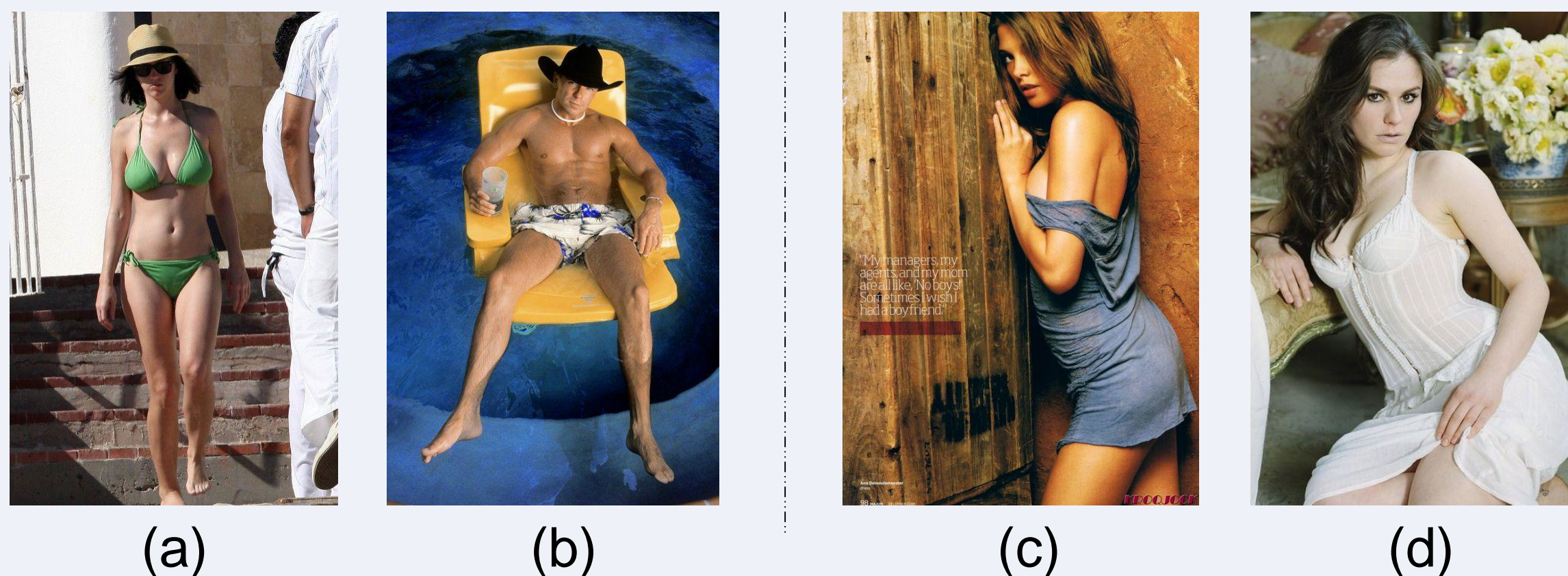
❑ Images (a) and (b) show subjects who are almost nude, but most humans will agree that these images contain no sexual intent.
❑ On the other hand, images (c) and (d) do not contain nude subjects, but they clearly show sexual intent.

Approach

Sexual Intent
Yes, *M*aybe, *N*o

Moods and Emotions
*D*efensive, *S*uggestive, *P*layful, *R*elaxed, *U*pset

Facial Expressions
*L*ooking, *E*yebrow, *S*mile, Eye li*d*s, *M*outh, *B*iting lips

Posture & Gesture
Body *P*osture,
Body Position and *M*ovement,
*B*ody Facing Camera,
*F*ace Facing Camera,
*H*ead Position,
*S*pread Eagles,
Behind Head *E*lbow Direction
Position of *W*rists,
*G*esture with Fingers

Scene Context
Outdoor *S*cene,
Outdoor *E*vent,
Indoor with *P*rops,
Indoor with *F*lat Back

Skin Exposure
Fully *C*lothed,
*B*are Bodied,
*P*rivate Parts Exposed

Automatic Features
Color Histogram, HOG, SIFT,
CaffeNet FC6, FC7, FC8

❑ Hierarchical Framework is a multi-layer learning model with each layer consisting of a set of multi-class SVMs.
❑ SVMs in first level are trained using automatic features to predict Attributes.
❑ SVMs in second level are trained using the 17 attributes to predict Moods and Emotions.
❑ SVMs in third level are trained using 5 moods and emotions to predict global Sexual Intent.

Dataset

❑ 1,146 celebrity images
  ➤ 203 Hollywood celebrities from *people.com*
  ➤ 892 and 254 images of female and male candidates respectively
  ➤ 5.6 images per person ratio
❑ 19 questions per image for annotations
  ➤ Amazon Mechanical Turk by majority voting of 3 annotators
  ➤ 70.5% annotator consensus
❑ Available online
  ➤ https://github.com/DebashisGanguly/SexualIntentDetection

Results

❑ We compare performance of Direct and Hierarchical models trained from different automatically extracted features.
  ➤ Direct: Single level of classification hierarchy
  ➤ Hierarchical: Multiple levels of multi-class classifiers
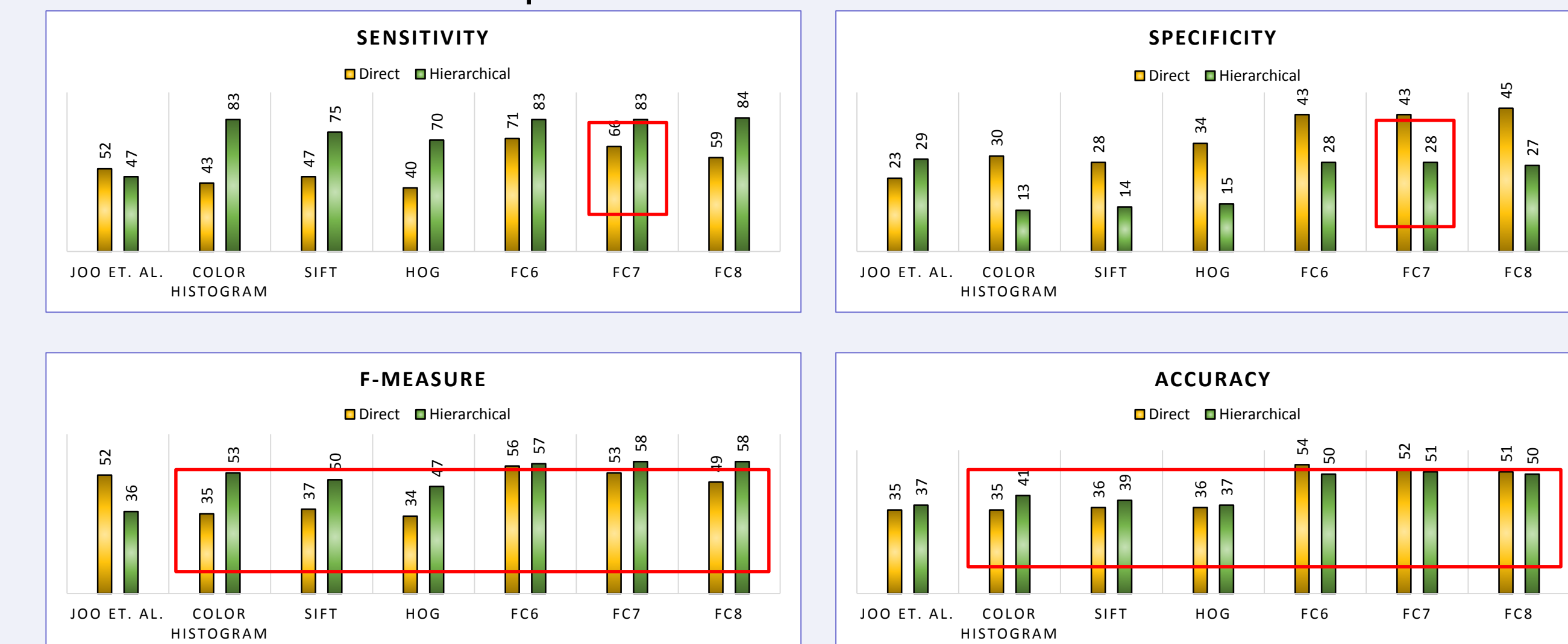
❑ F-measure and accuracy are similar for both methods.
❑ Hierarchical framework has higher sensitivity compared to corresponding Direct model.
  ➤ In a real world application, higher sensitivity imposes the ability to catch any and all sexually provocative images.
❑ Hierarchical framework has lower specificity compared to corresponding Direct model.
  ➤ Lower specificity means higher false positive rate. This results in classifying non-provocative images as sexually provocative images.
  ➤ This means some images will need to be manually checked, but users are accustomed to this review latency.

Conclusion

❑ Our method enables automated contents classification based on behaviors and intents of the portrayed subjects.
❑ It allows prompt intervention of human experts upon integrating the proposed methodology with mobile apps, social media websites, and media streaming websites.