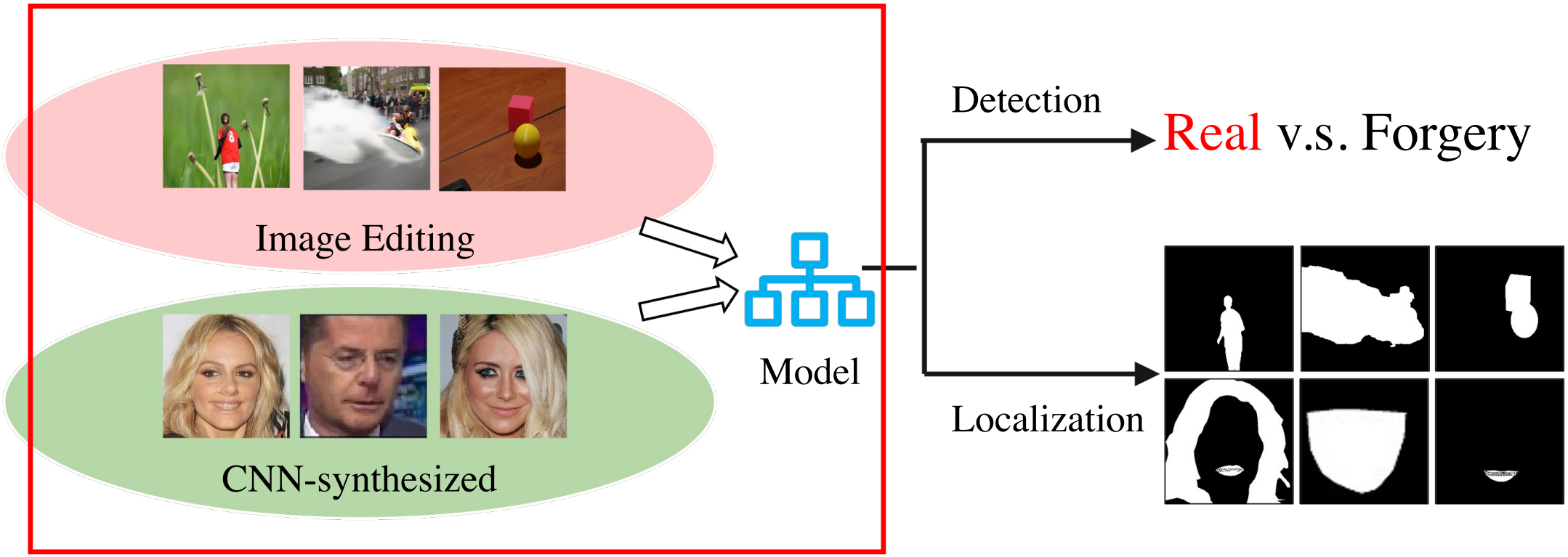
Hierarchical Fine-Grained Image Forgery Detection and Localization

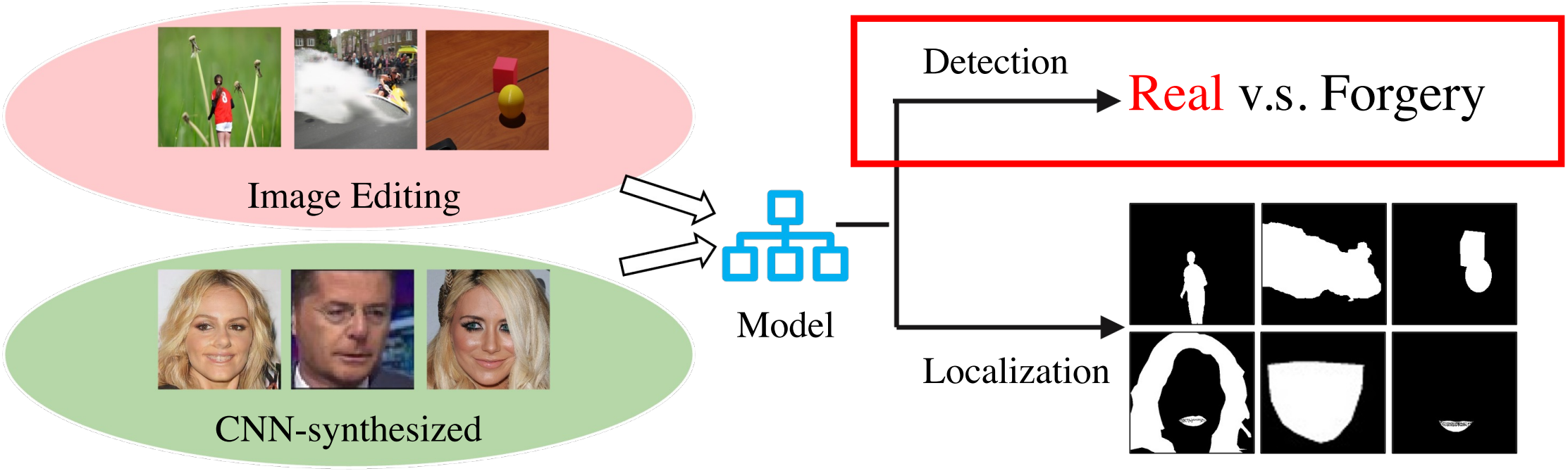
Xiao Guo, Xiaohong Liu, Zhiyuan Ren, Steven Grosz, Iacopo Masi, Xiaoming Liu
[1]Michigan State University, [2]Shanghai Jiao Tong University, [3]Sapienza University of Roma

Poster ID: 300
Session: TUE-AM-300

Introduction



Introduction



Introduction

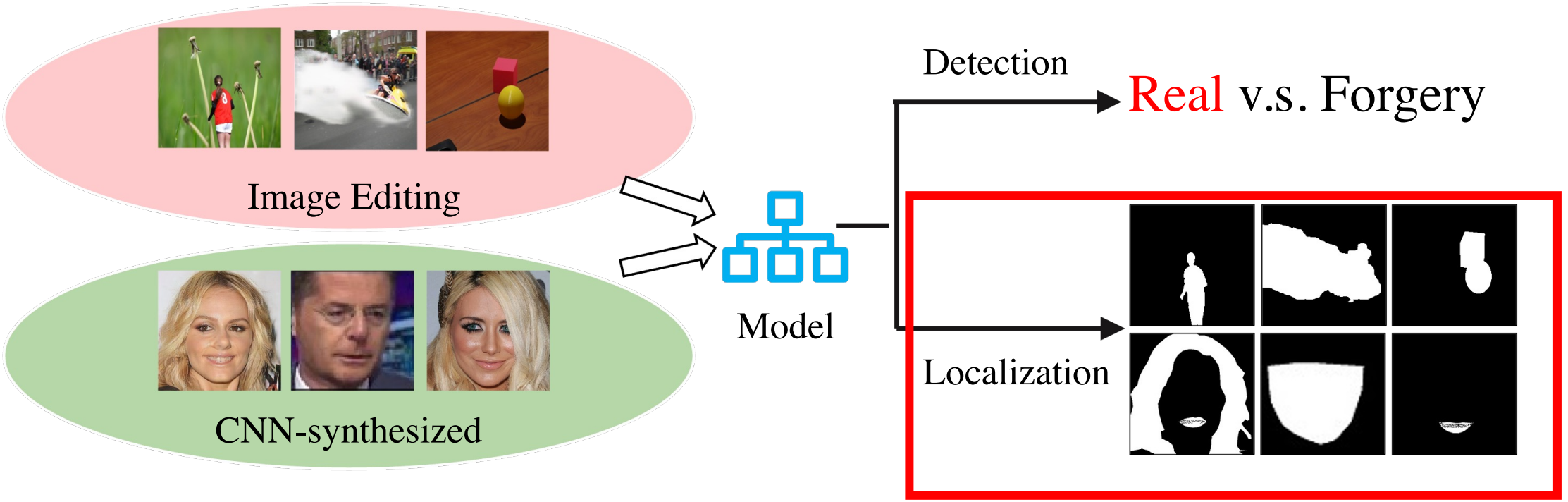
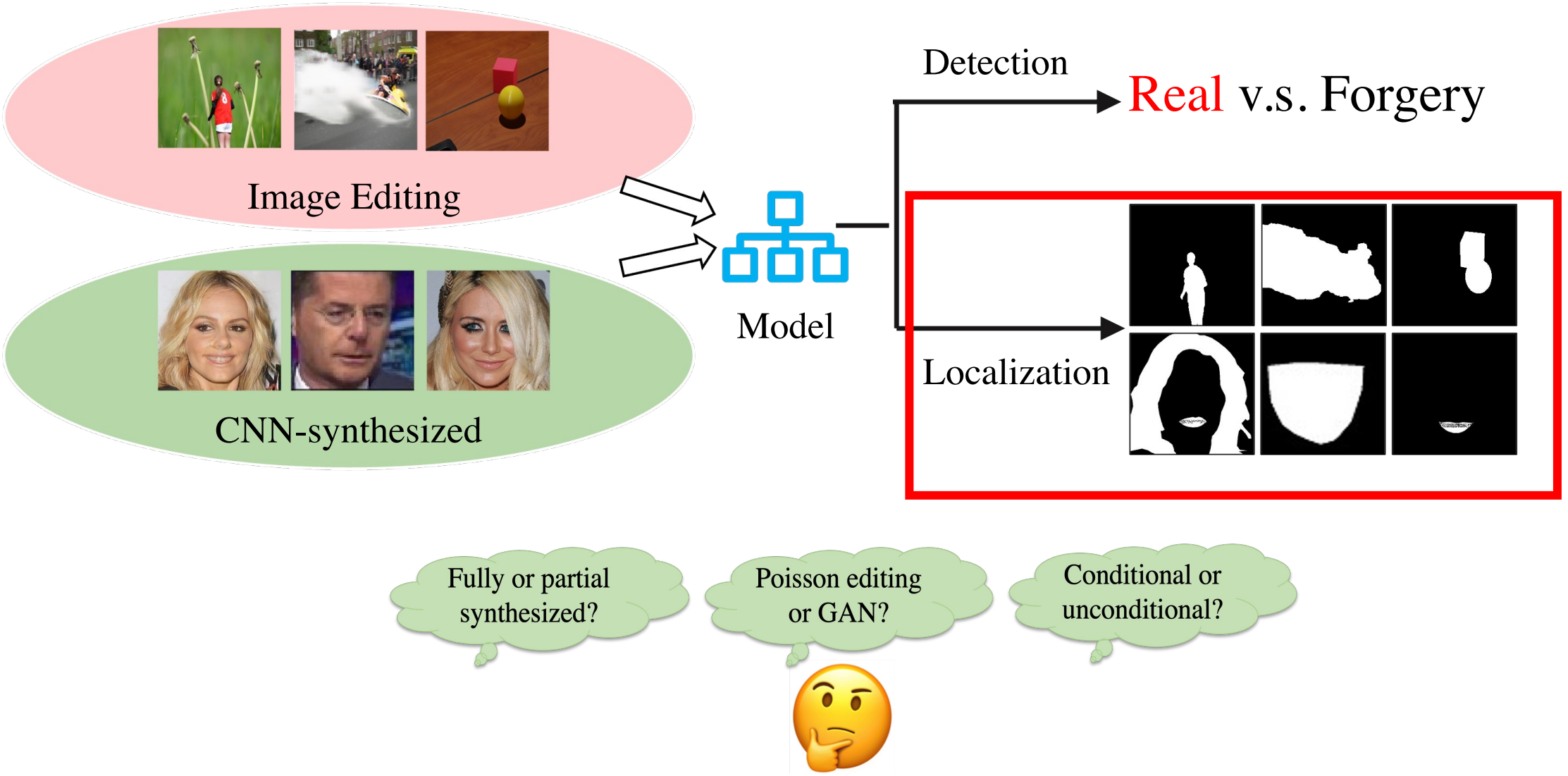


Image Forgery Detection and Localization (IFDL)



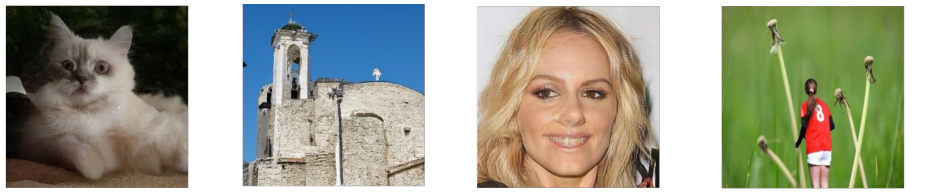
Brief Summary

In this work, we have three contributions.

- A hierarchical fine-grained formulation helps learn comprehensive representation for IFDL.
- We propose HiFi-Net, which performs well on IFDL, also identifies a diverse spectrum of forgery attributes.
- A new dataset (HiFi-IFDL) to facilitate the hierarchical fine-grained IFDL study.

Hierarchical Fine-grained Formulation

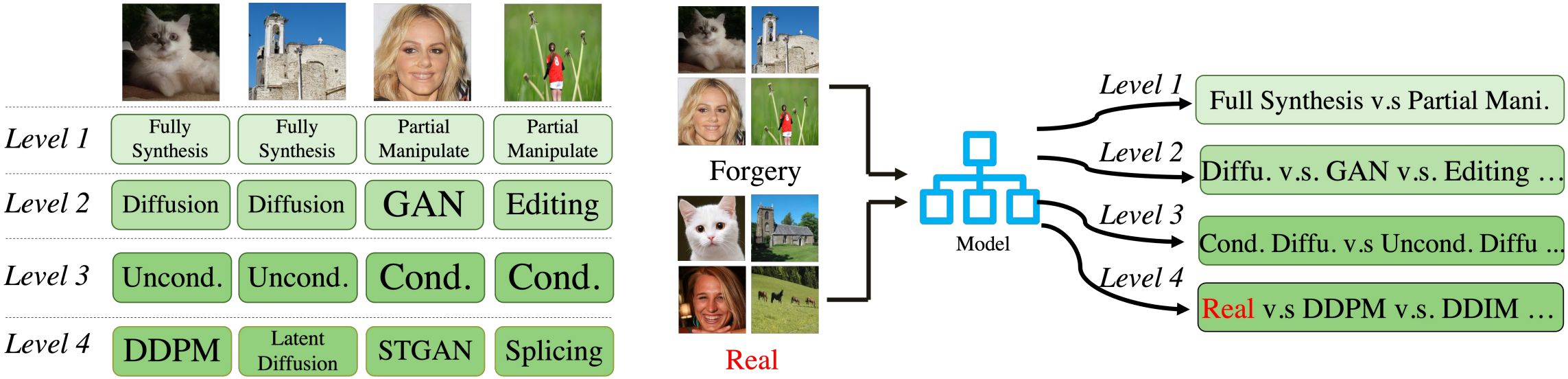
We first interpret the forgery attribute of each image



Level 1	Fully Synthesis	Fully Synthesis	Partial Manipulate	Partial Manipulate
Level 2	Diffusion	Diffusion	GAN	Editing
Level 3	Uncond.	Uncond.	Cond.	Cond.
Level 4	DDPM	Latent Diffusion	STGAN	Splicing

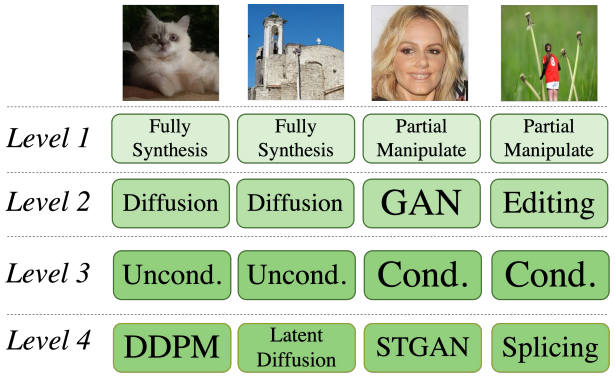
Hierarchical Fine-grained Formulation

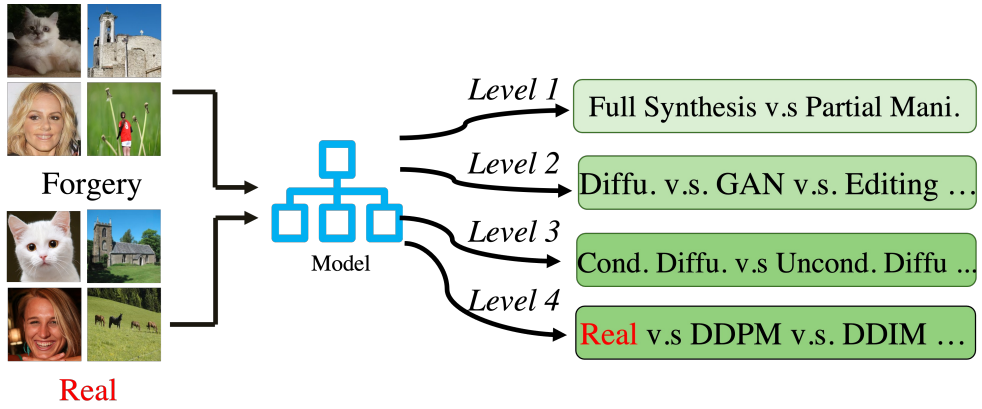
We encourage to learn multi-level forgery attributes.

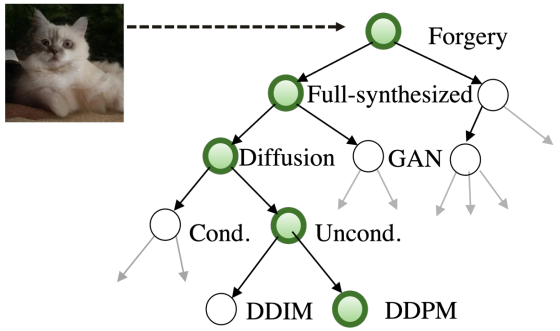


Hierarchical Fine-grained Formulation

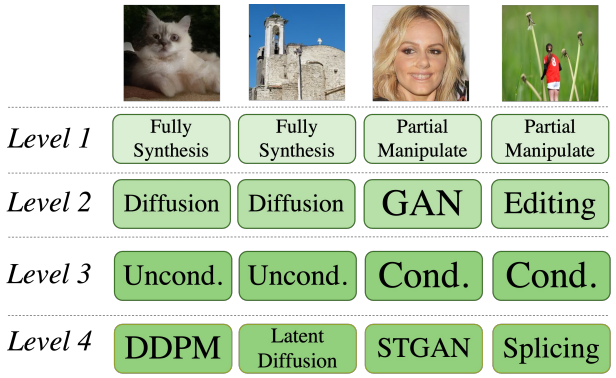
Following the hierarchical nature.

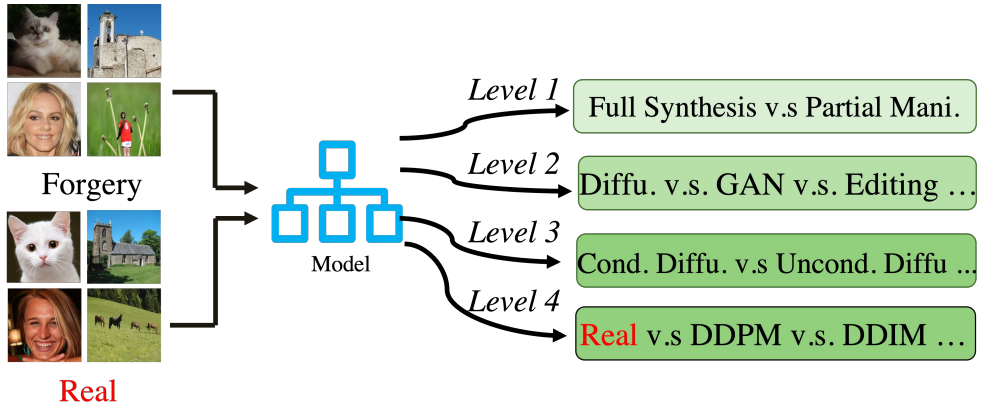


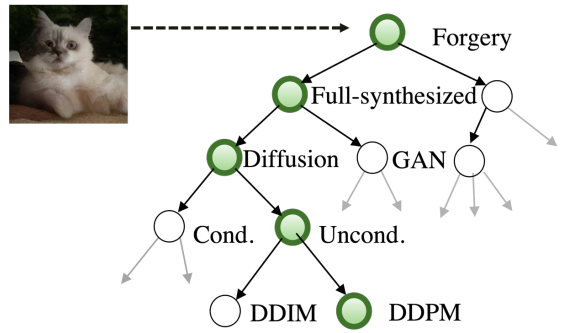


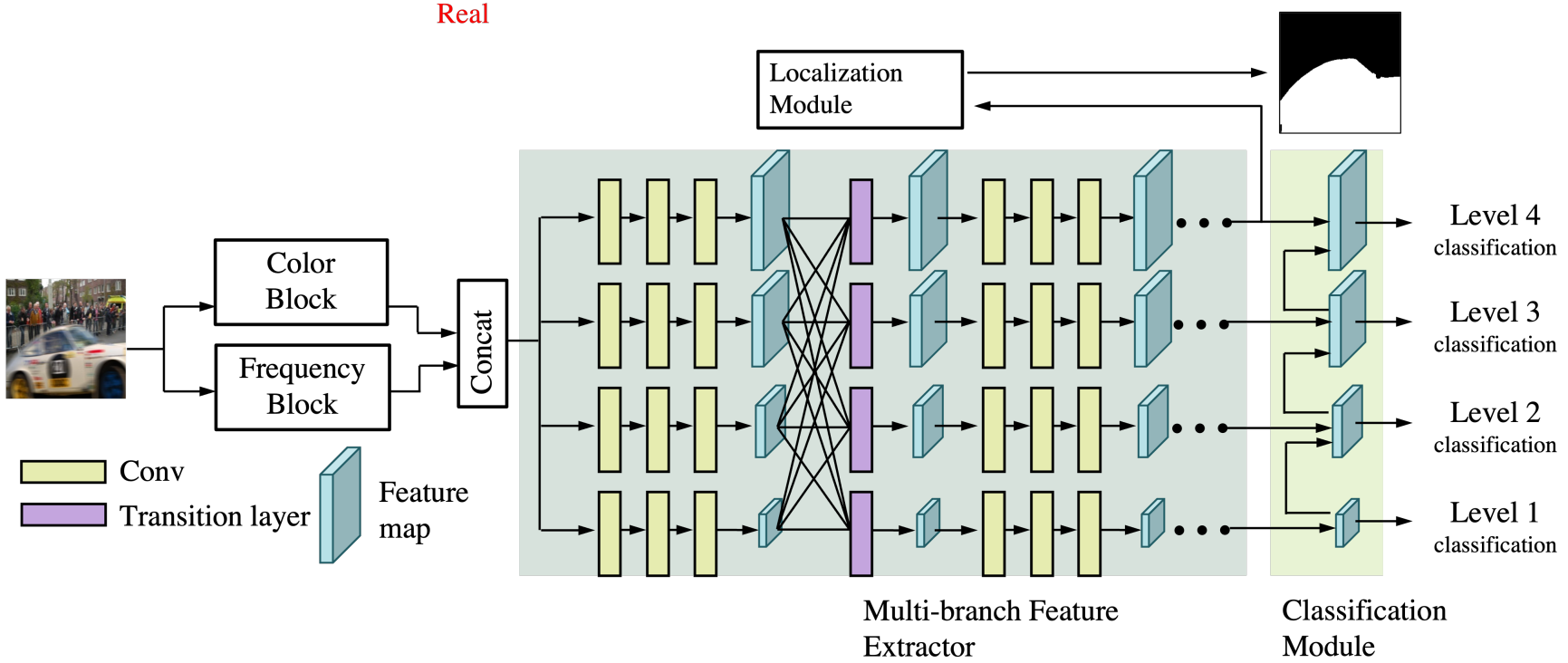


Hierarchical Fine-grained Network (HiFi-Net)

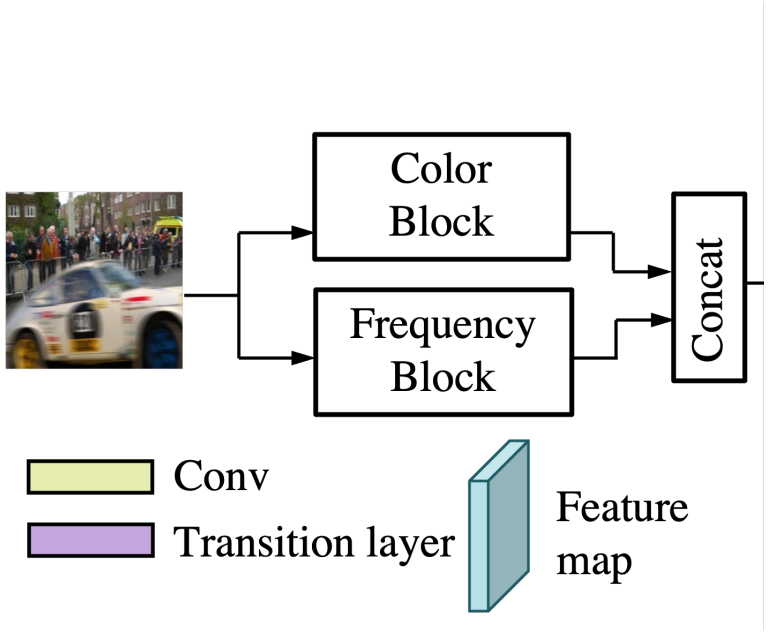




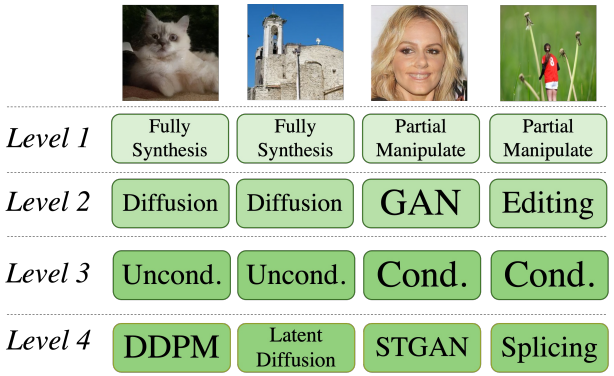


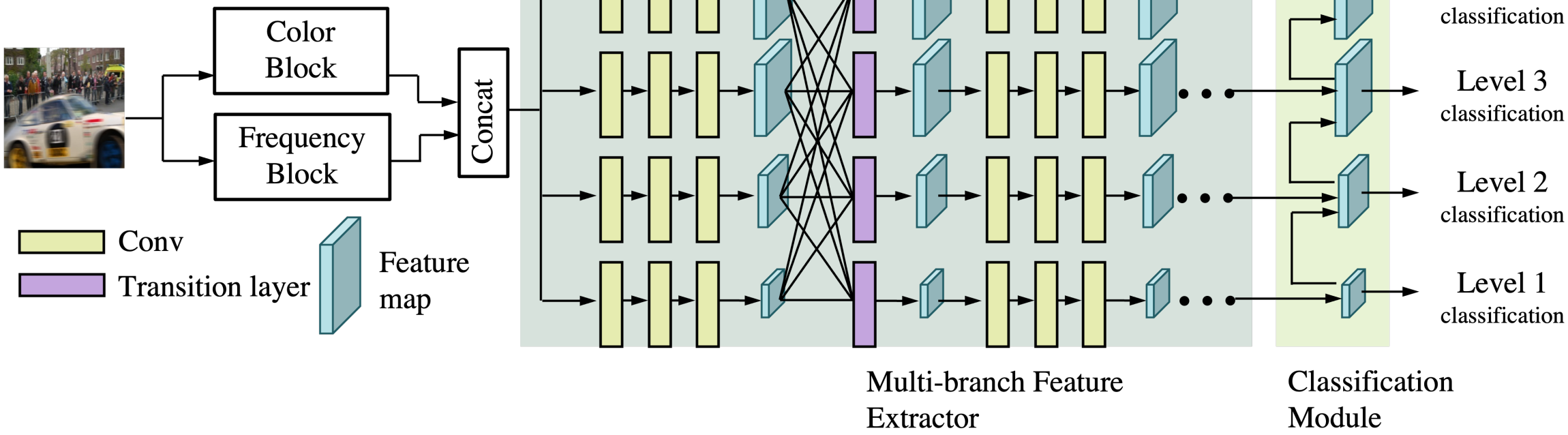


Hierarchical Fine-grained Network (HiFi-Net)

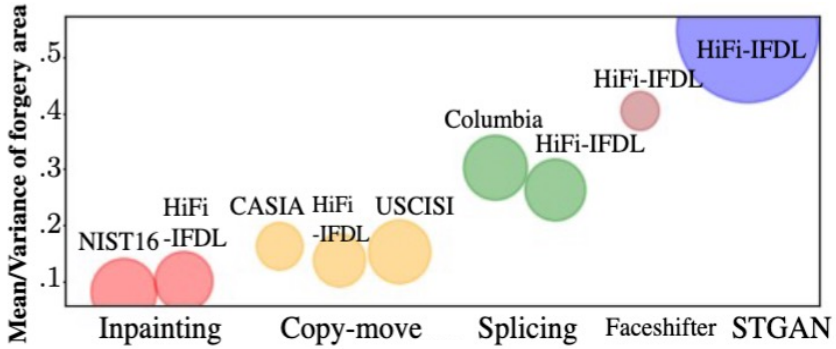


Hierarchical Fine-grained Network (HiFi-Net)





Hierarchical Fine-grained Network (HiFi-Net)



Color Block

Frequency Block

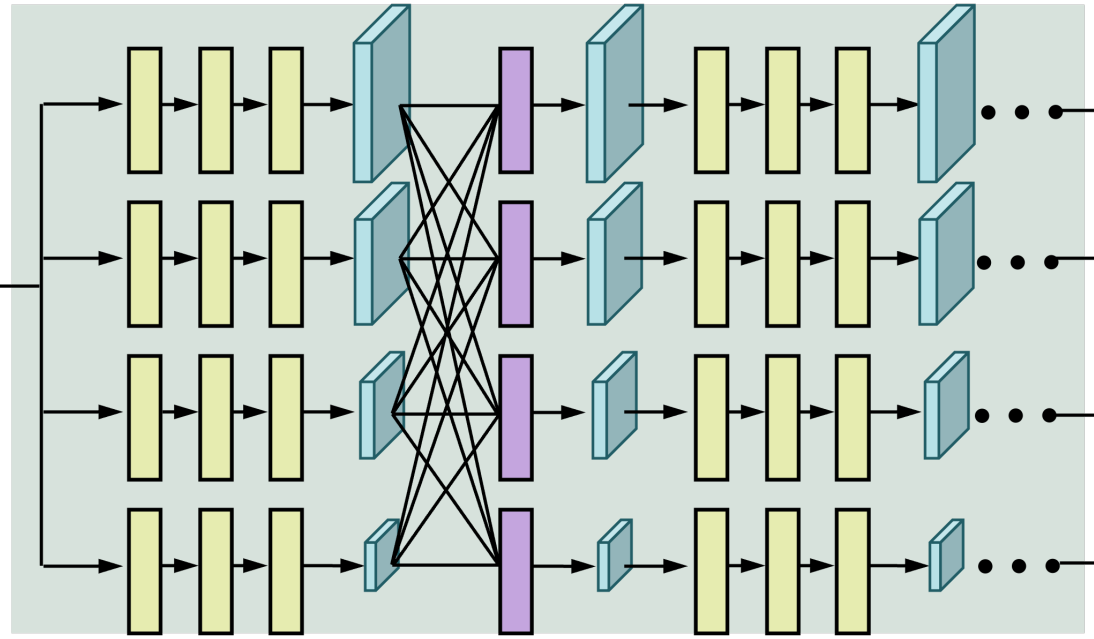
Concat

Conv

Transition layer

Feature map

Localization Module



Level 4 classification

Level 3 classification

Level 2 classification

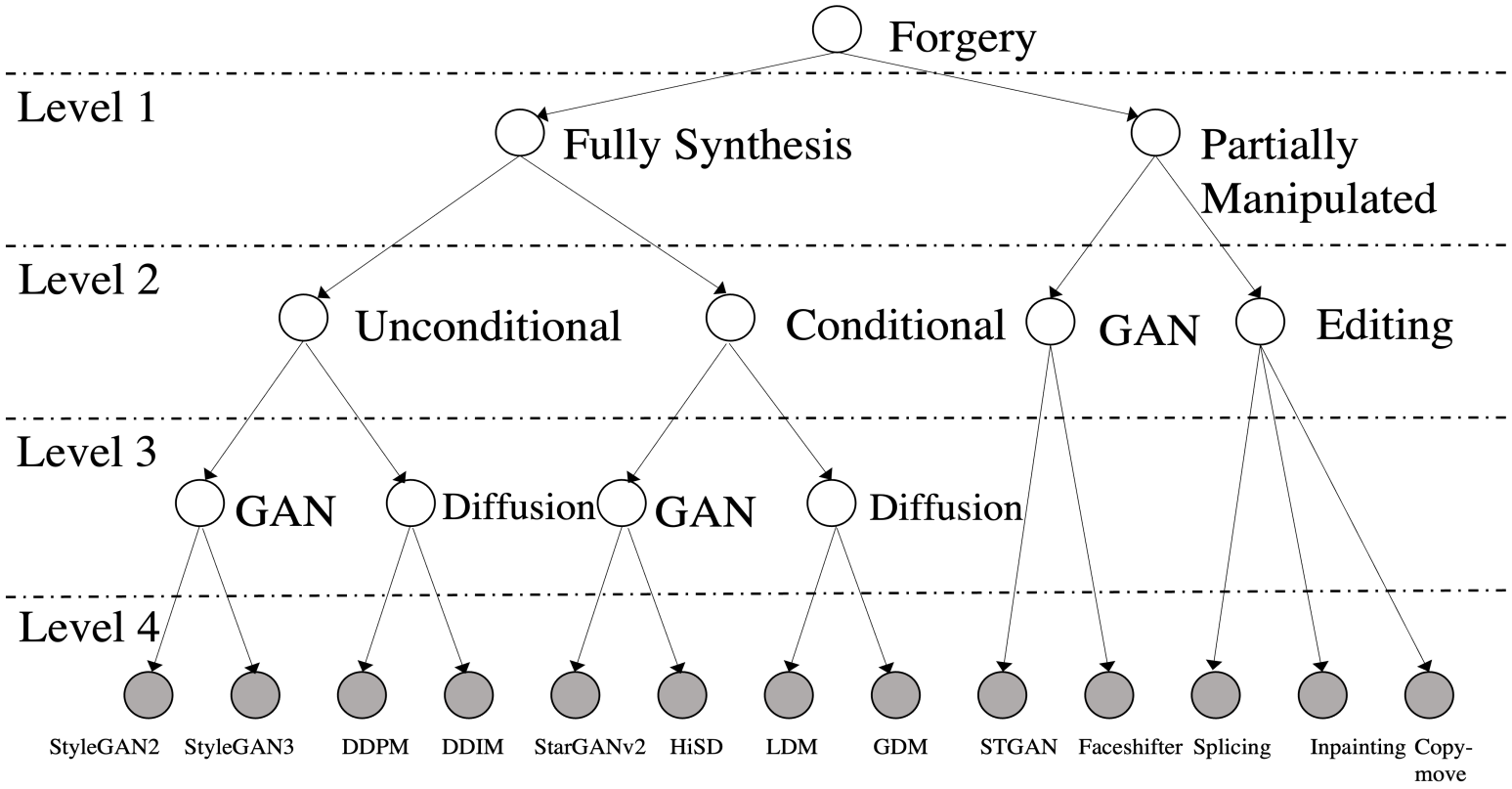
Level 1 classification

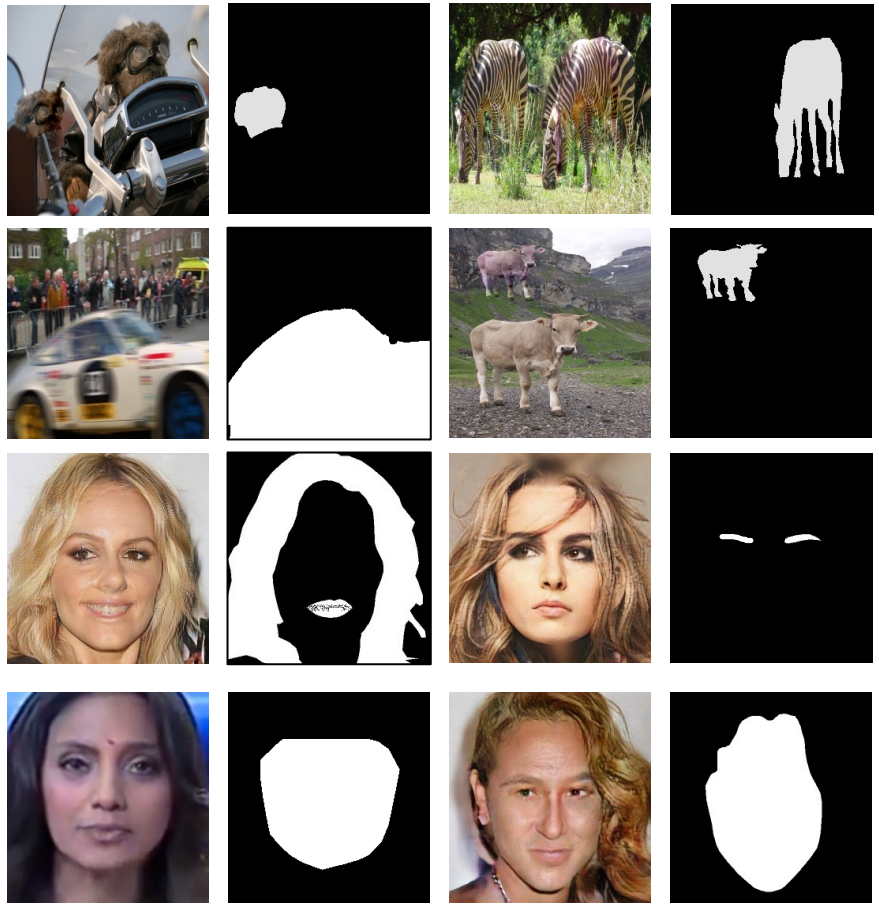
Multi-branch Feature Extractor

Classification Module

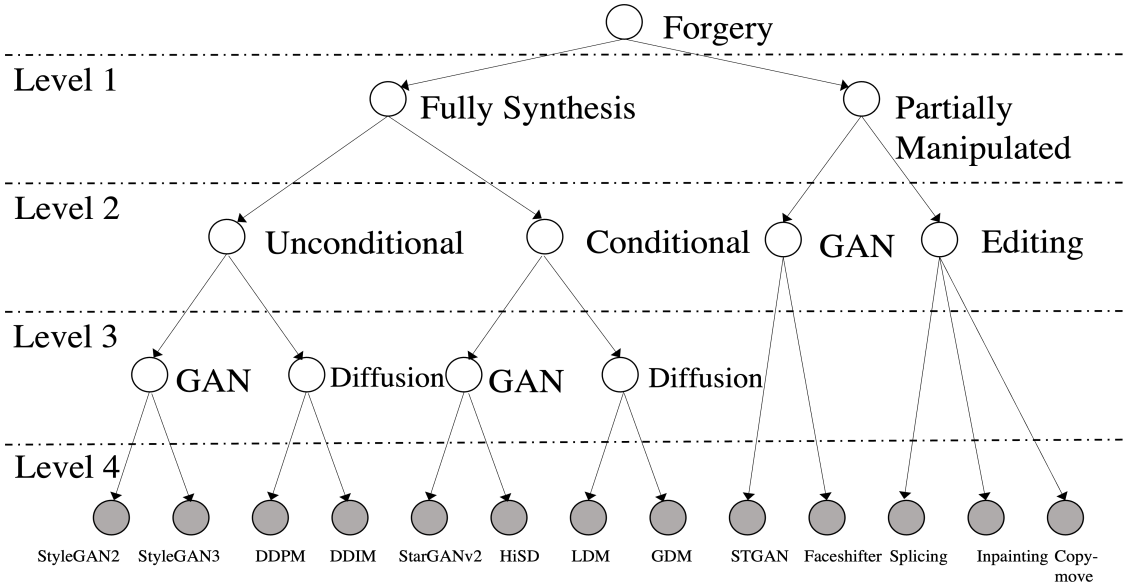
HiFi IFDL dataset

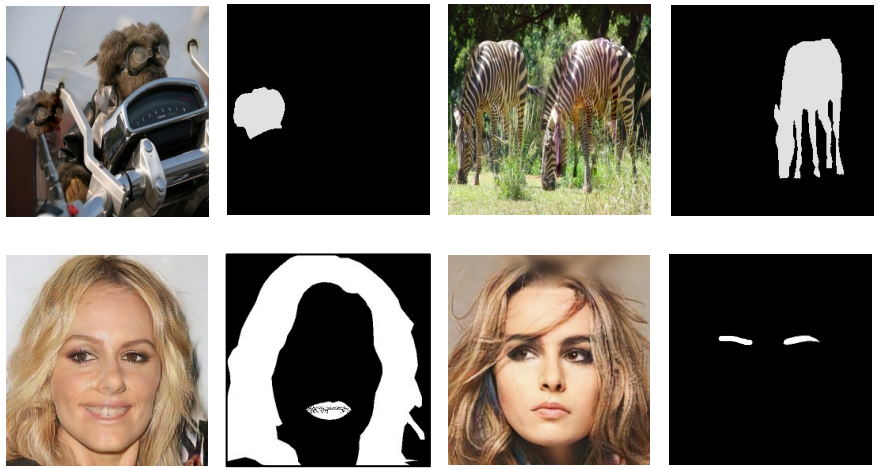
We construct a dataset to facilitate HiFi formulation.

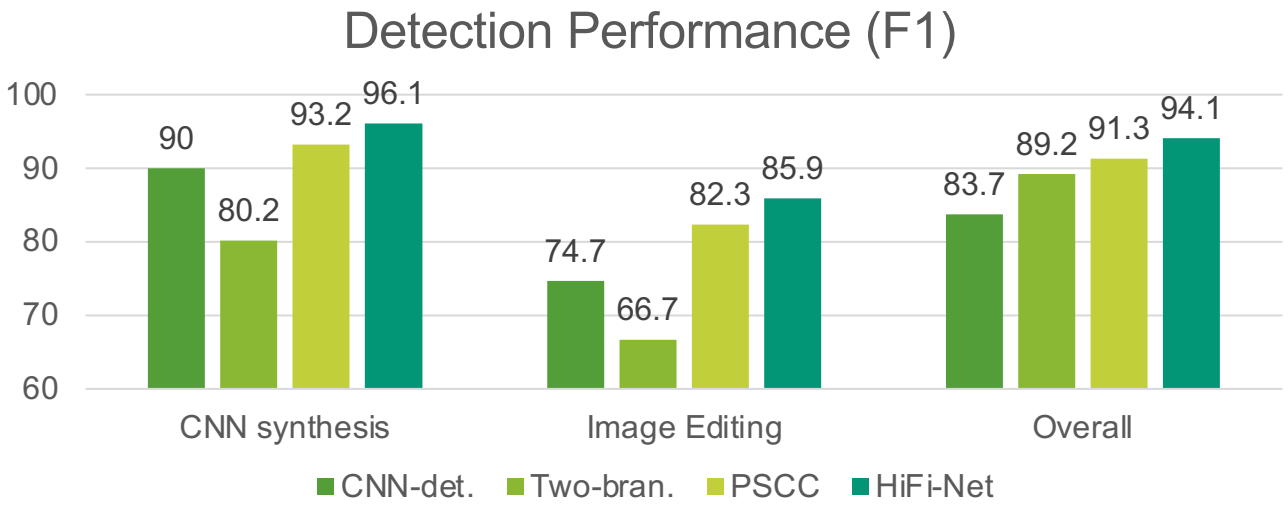


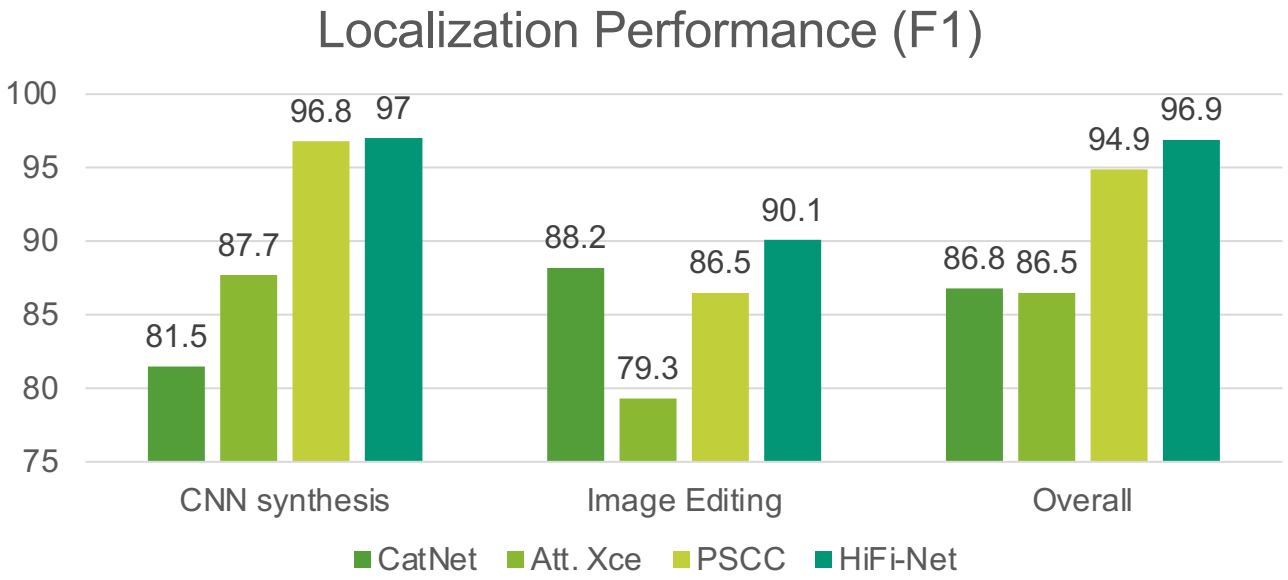


Detection and Localization performance



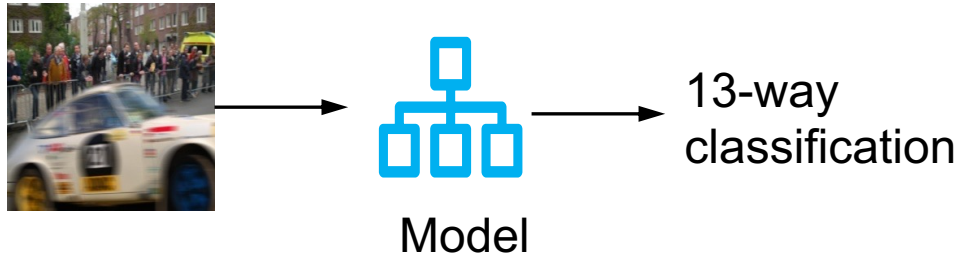


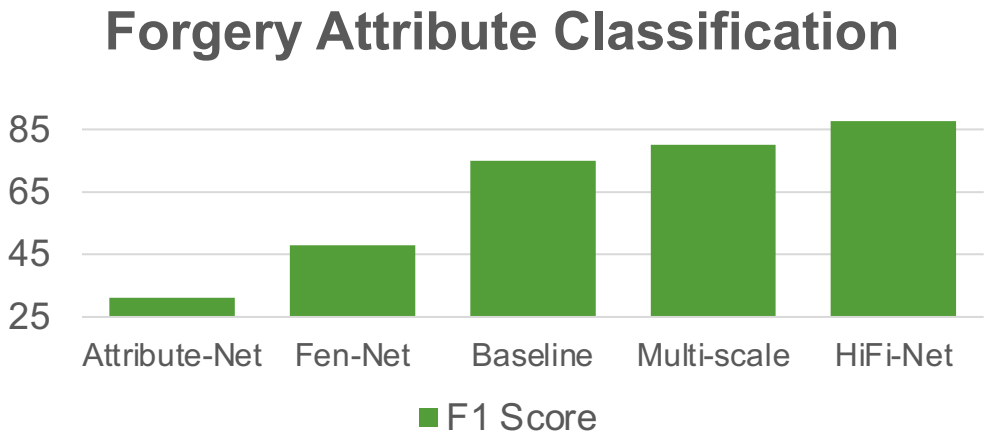


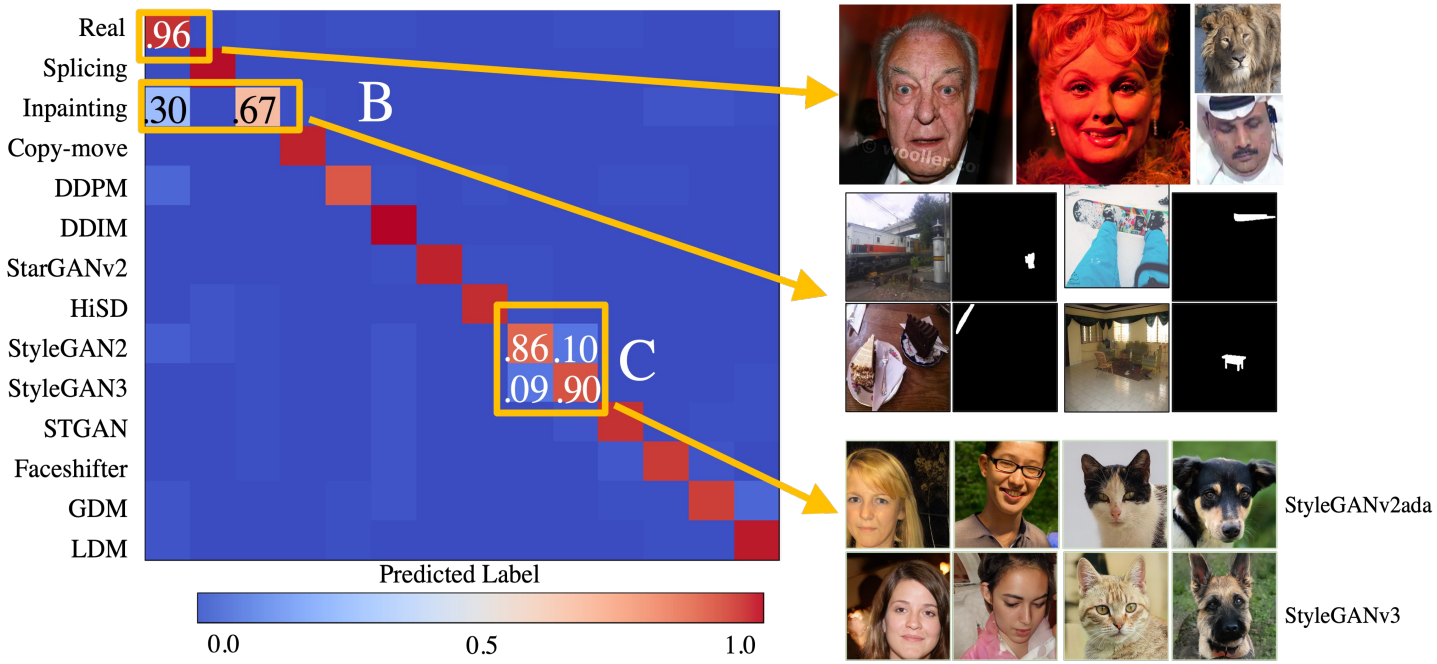


Forgery Attribute Classification

Forgery Attribute learns what forgery method used to generate the input image.

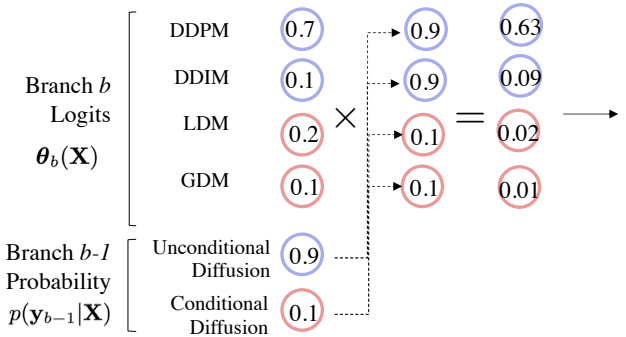


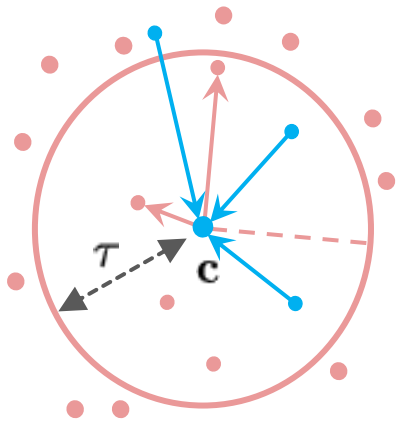




More Visualizations

More analysis on the classification and localization can be found in the poster.

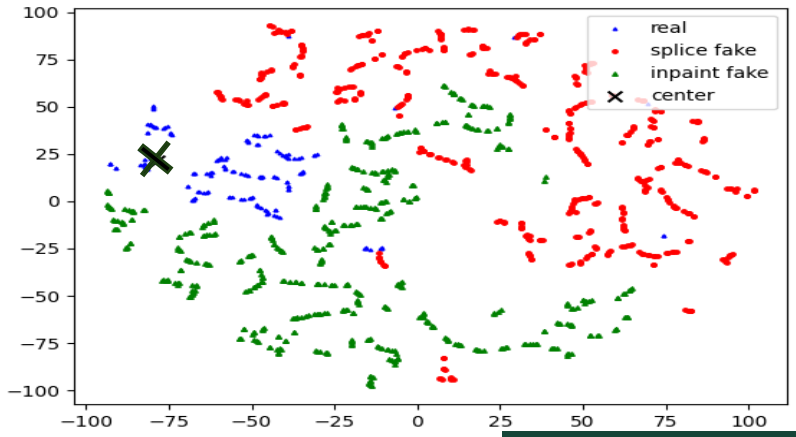




$$\mathcal{L} = \begin{cases} \left\| \mathbf{F}'_{ij} - \mathbf{c} \right\|_2 & \text{if } \mathbf{M}_{ij} \text{ real} \\ \max\left(0, \tau - \left\| \mathbf{F}'_{ij} - \mathbf{c} \right\|_2\right) & \text{if } \mathbf{M}_{ij} \text{ forged.} \end{cases}$$

$\mathbf{M}_{ij}$ is the input pixel

$\mathbf{F}_{ij}$ is the input pixel feature



Thank You!