Detecting Photorealistic Computer Graphics using Convolutional Neural Networks
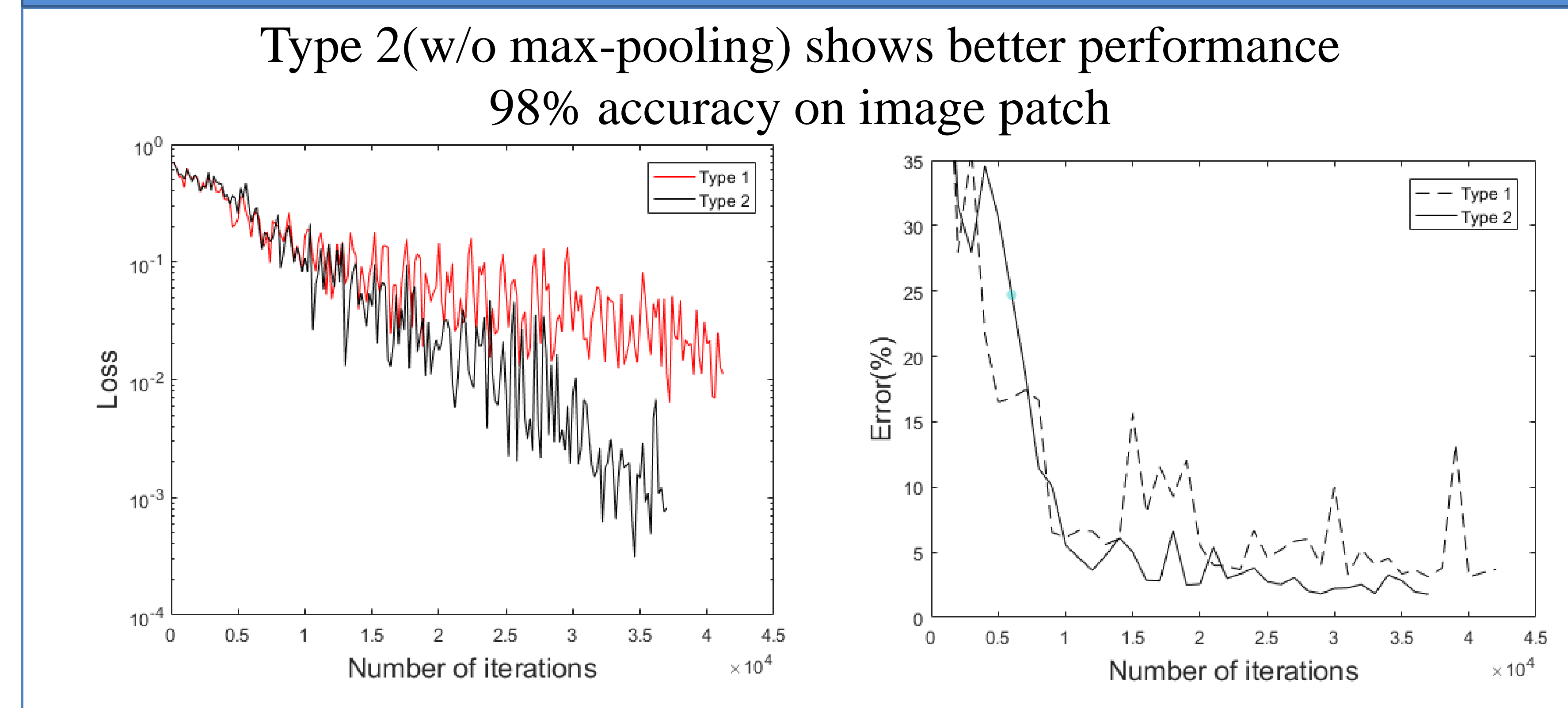
In-Jae Yu, Jin-Suk Park, Do-Guk Kim, Jong-Uk Hou, Sung-Hee Choi, Heung-Kyu Lee

Department of Computer Science, KAIST, 373-1,
Guseong-Dong, Yuseong-gu, Daejeon, Republic of Korea

{ijyu@mmc.kaist.ac.kr, jspark@mmc.kaist.ac.kr, dgkim@mmc.kaist.ac.kr, juheo@mmc.kaist.ac.kr, sunghee@kaist.edu, hklee@mmc.kaist.ac.kr}

Research objectives - With the development of rendering and GPU technology, the artificial modeling of the scene has become more sophisticated and the results has became more realistic after rendering. Our goal is distinguishing these scene with real photograph.

Limitation of previous method – 1)Existing PRCG detection methods use various handcrafted features, but there are no clear relationship among them. 2)Method using entire image can't handle the case when photograph and PRCG are synthesized into one image.
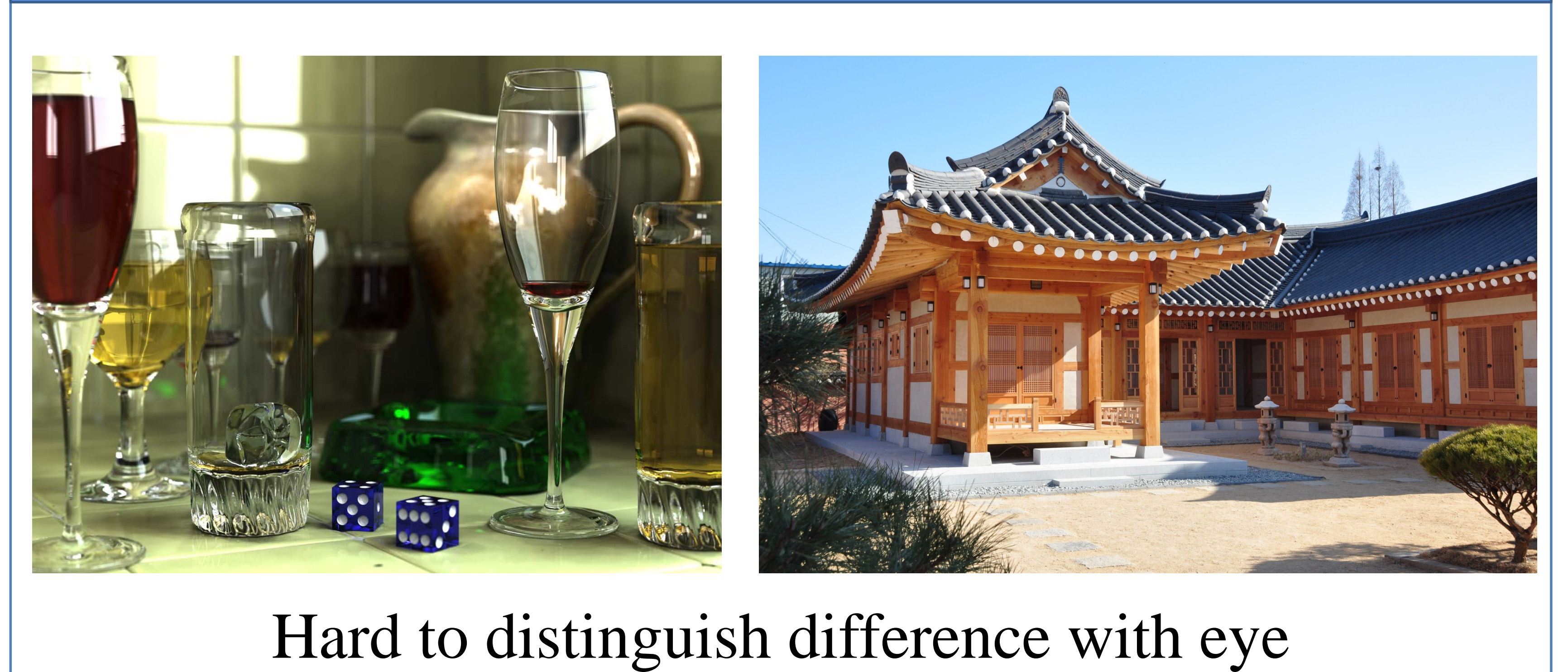
Proposed Method - We trained a network to classify input image patches to PRCG or photograph. Using the trained network, we explored the ratio/distribution of PRCG patches. We trained two VGG –like network with/without max-pooling.

Experiment Results – Network training result shows that with simple network architecture, it is possible to classify whether the small patch is from Photograph or PRCG. Also, result on whole image shows that this method can classify the image to PRCG/Photograph or find suspicious synthesized region.
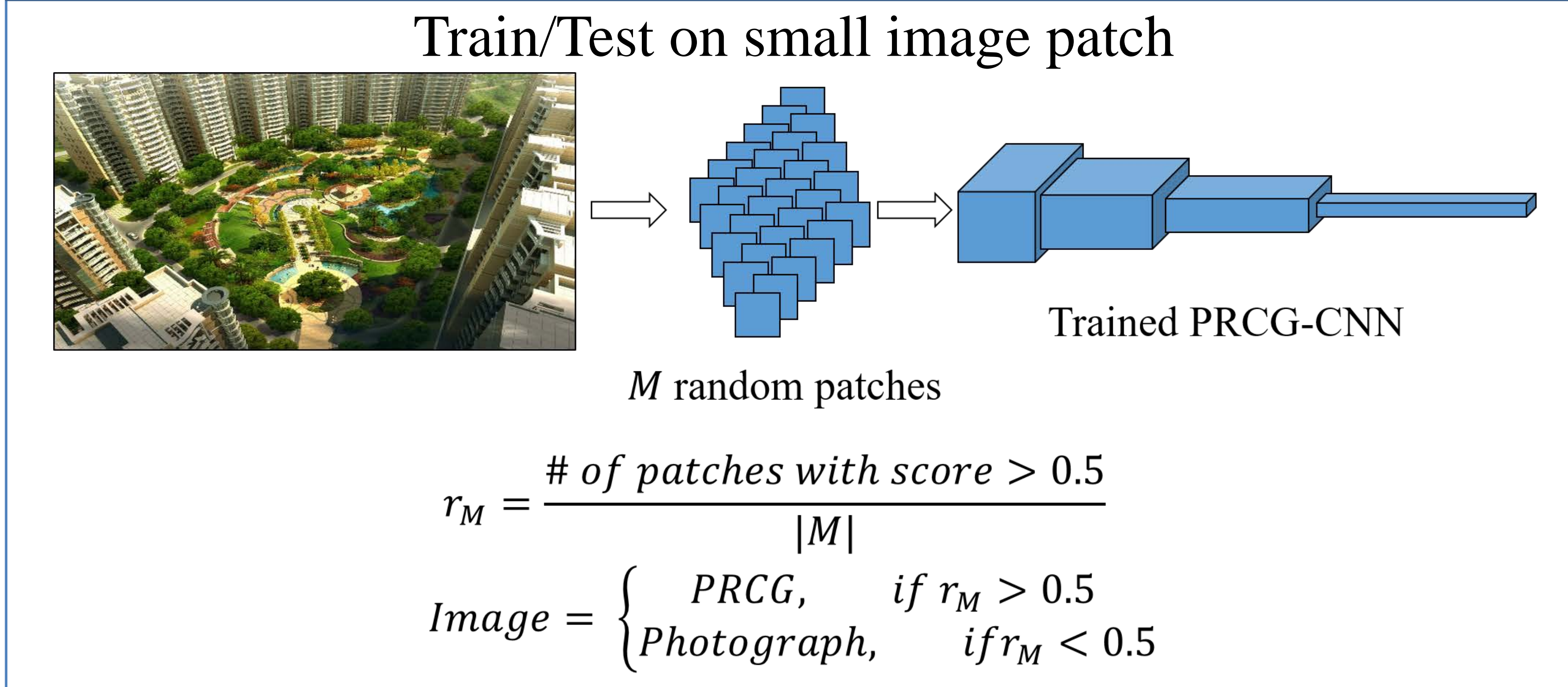
Photorealsitc Computer Graphics(PRCG)
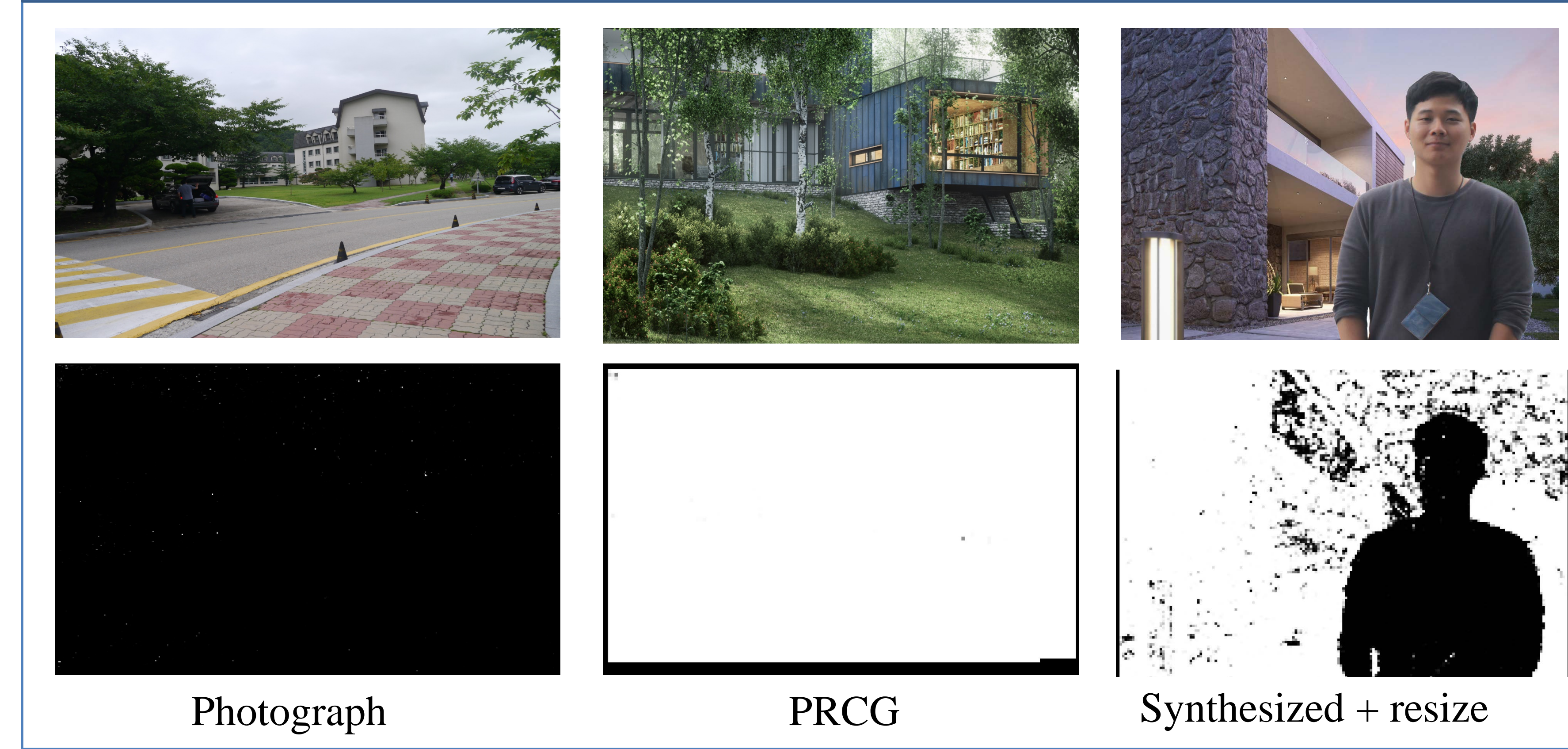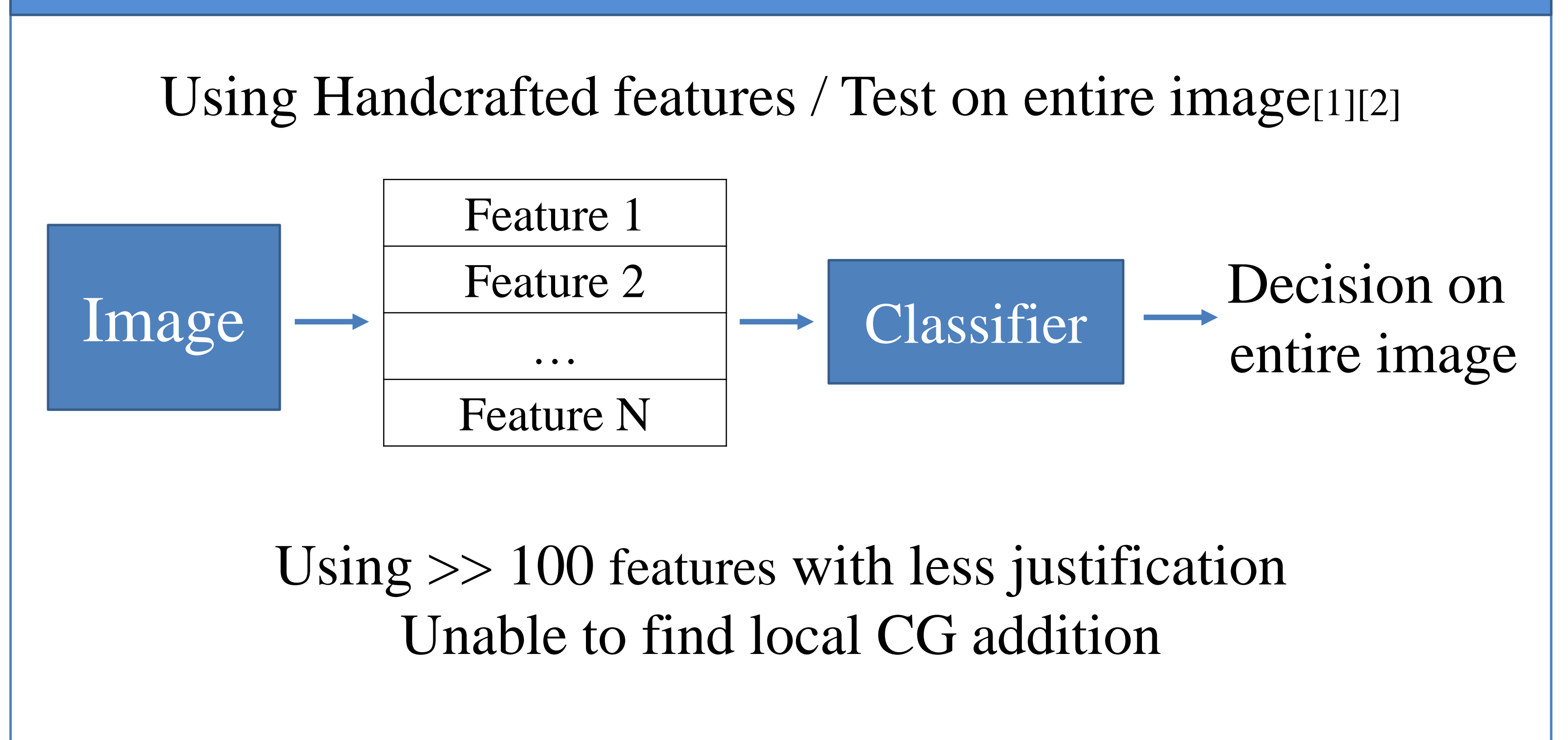

Hard to distinguish difference with eye

Previous Method

Using Handcrafted features / Test on entire image[1][2]

Using >> 100 features with less justification
Unable to find local CG addition

Proposed Method

Train/Test on small image patch

M random patches

Trained PRCG-CNN

$$r_M = \frac{\# \text{ of patches with score} > 0.5}{|M|}$$

$$Image = \begin{cases} PRCG, & if\ r_M > 0.5 \\ Photograph, & if\ r_M < 0.5 \end{cases}$$

Network Architecture

VGG Style **with/without** max-pooling

Type 1	Conv 32	Max pooling	Conv 64	Max pooling	Conv 128	Max pooling	FC 1024	FC 2
	Conv 32		Conv 64		Conv 128		FC 1024	Softmax
Type 2	Conv 32	Conv 64	Conv 128	FC 1024	FC 2			
	Conv 32	Conv 64	Conv 128	FC 1024	Softmax			

Network Training Results

Type 2(w/o max-pooling) shows better performance
98% accuracy on image patch

Experimental Results

| Photograph | PRCG | Synthesized + resize |

Conclusion

We used simple VGG style network for detecting photorealistic computer graphics. Experiment result shows that our method shows good result on entire image, and also for detecting synthesized regions.

Unlike other computer vision problems, forensics deal with very low dimensional features. Our future work is finding network architecture that fits to digital forensics

[1] Tian-Tsong Ng and Shih-Fu Chang, "Classifying photographic and photorealistic computer graphic images using natural image statistics," 2004.
[2] Siwei Lyu and Hany Farid, "How realistic is photorealistic?," IEEE Transactions on Signal Processing, vol. 53, no. 2, pp. 845–850, 2005.