Fast Object Detection Algorithm using Edge-based Operation Skip Scheme with Viola-Jones Method

Cheol-Ho Choi*, Joonhwan Han, Jeongwoo Cha, Jungho Shin, and Hyun Woo Oh
cheoro1994@hanwha.com

Core H/W Team, Pangyo R&D Center
Hanwha Systems Company, Ltd.

2024 IEEE International Conference on Artificial Intelligence Circuits and Systems

1. Motivation

Motivation (1/2)

- **European New Car Assessment Programme (Euro-NCAP)**[1]
  - Child presence detection system (CPD, after 2025)
    - Direct sensing (ex. Pressure sensor)
    - Indirect sensing (ex. Camera sensor)
  - To meet the requirements, an object detection algorithm must be included in the product
    - Low-cost for mass production
    - Reasonable detection performance

Fig 1. Euro ENCAP Standard

[1] https://www.euroncap.com/media/79888/euro-ncap-cpd-test-and-assessment-protocol-v12.pdf

Motivation (2/2)

- **Algorithm Selection**
  - Low-cost
    - A small number of parameters for classification (for memory capacity)
    - Fewer classification operations (for processor cost)
  - Fast processing speed
  - Reasonable detection performance

  →Among machine learning algorithms, we select Viola-Jones method
    - Haar-like features
    - Cascade classifier architecture

2. Background

Background (1/3)

- **Haar-like Rectangle Features**[2]
  - Suitable for machine learning classifier models
    - Cascade Classifier
    - SVM (support vector machine)
  - Fast processing speed performance
    - Summed-area table (integral image)
  - Reasonable detection performance
    - Brightness difference for target-oriented train feature generation

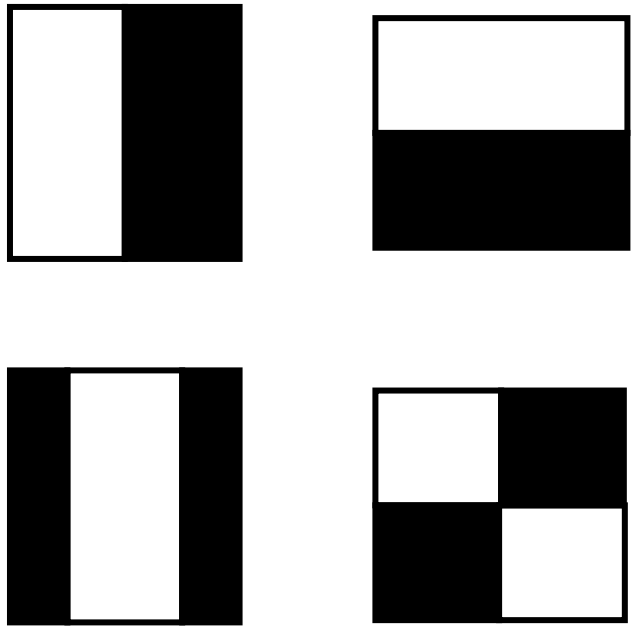


Fig 2. Rectangle Haar-like feature examples

[2] P. Viola and M. Jones, "Rapid Object Detection using a Boosted Cascade of Simple Features," *in Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, 2001.

Background (2/3)

- **Cascade Classifier Architecture[3] (1/2)**
  - Consists of multiple strong classifiers for step-by-step classification process
    - A strong classifier consists of trained Haar-like features
    - When the window satisfies all the strong classifier threshold values → True
    - When the window is not satisfies all the strong classifier threshold values → False

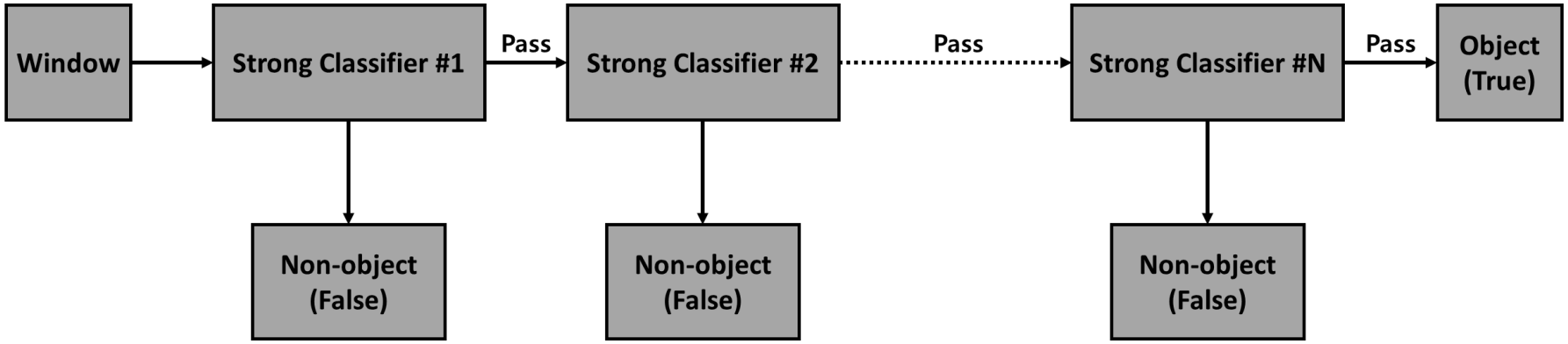


Fig 3. Cascade Classifier Architecture

[3] P. Viola and M. Jones, "Robust Real-time Face Detection," *International Journal of Computer Vision*, 56, 2024.

Background (3/3)

- **Cascade Classifier Architecture[3] (2/2)**
  - Advantage
    - Pipeline and parallel computation architectures can be designed
    - Reasonable detection performance by using step-by-step classification process
  - Disadvantage (Processing speed performance degradation)
    - Processing time increases when the input image resolution is larger
    - Processing time increases when the number of classification iteration is increased

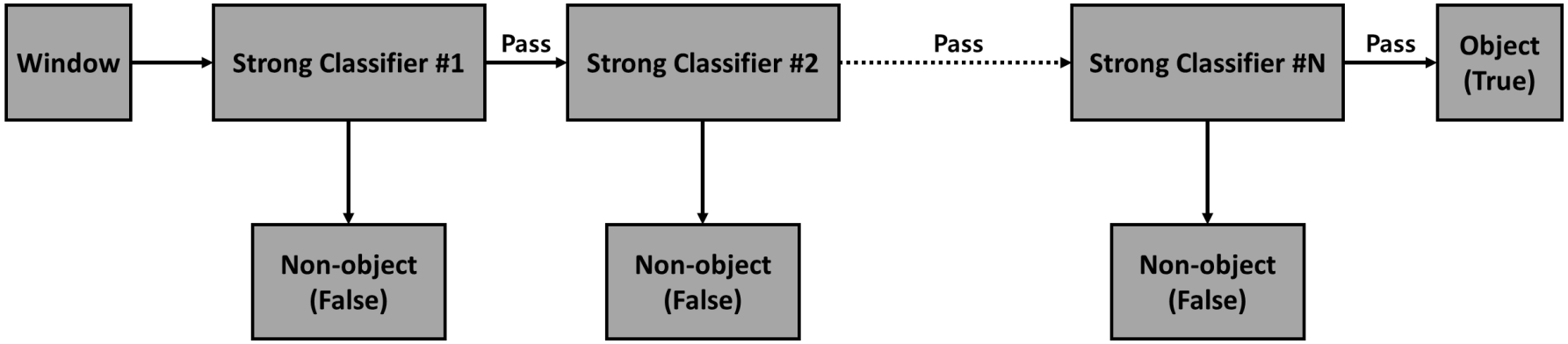


Fig 3. Cascade Classifier Architecture

3. Proposed Algorithm

Proposed Algorithm (1/3)

- **Edge Component Calibration**
  - Introduced the calibration concept[4]
  - 2-D Haar wavelet transform is used
    - Reduce the original input image resolution
    - Generate the image pyramid
  - Rectified Linear Unit (ReLU)
    - Extract the positive edge component values (horizontal, vertical, and diagonal directions)
  - Merged edge computation
    - $I_M = I_{R.H} + I_{R.V} - I_{R.D}$
  - Edge component calibration
    - $I_O = I_A + I_M$

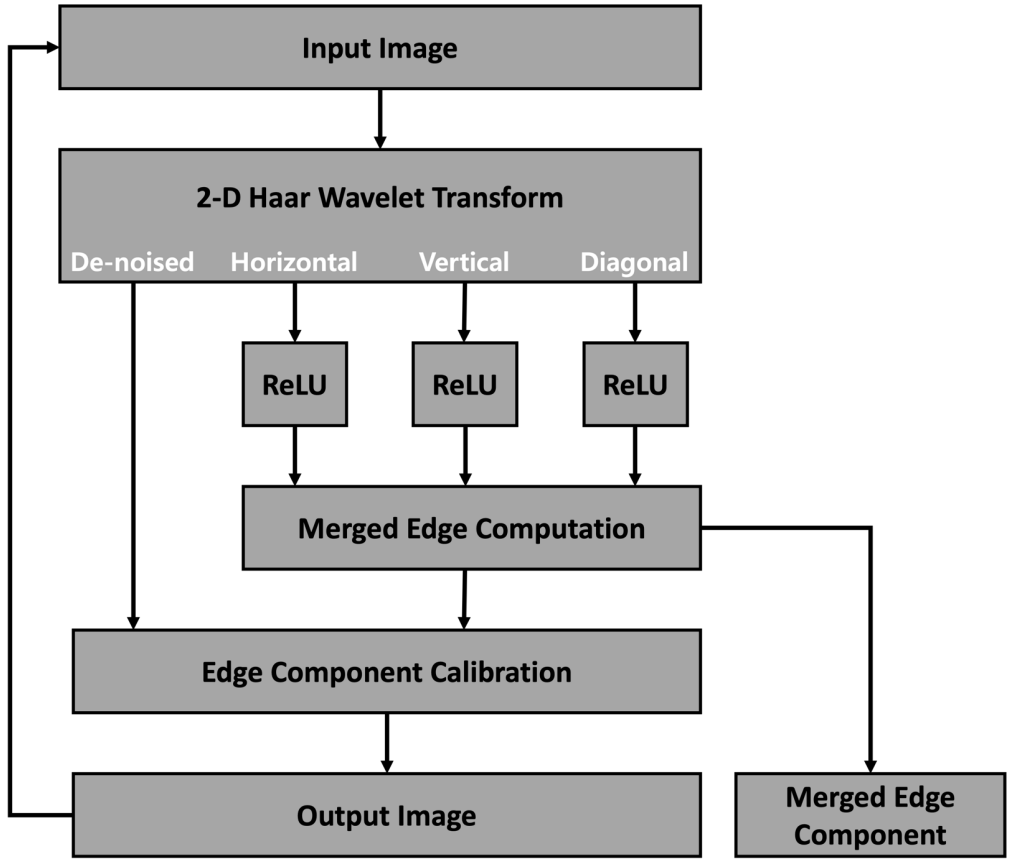


Fig 4. Proposed edge component calibration process

[4] C.-H. Choi, et al., "Face detection using haar cascade classifiers based on vertical component calibration," *Human-centric Computing and Information Sciences (HCIS)*, **12**(11), 2022.

Proposed Algorithm (2/3)

- **Edge-based Operation Skip (1/2)**
  - Reference coordinate is selected by using hyper-parameters
    - α : factor for selecting the vertical direction coordinate
    - β : factor for selecting the horizontal direction coordinate
  - Based on reference coordinate, extract the reference edge component value

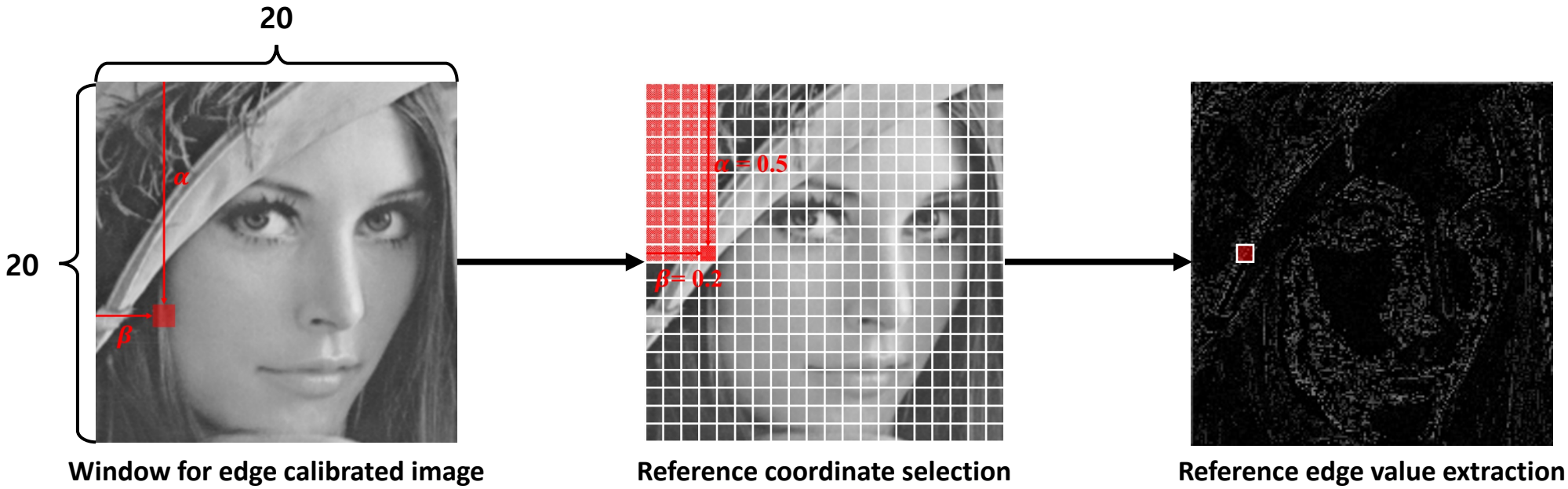


Fig 5. Proposed edge-based operation skip scheme process

Proposed Algorithm (3/3)

- **Edge-based Operation Skip (2/2)**
  - Operation mode
    - When the reference edge component value is positive
  - By-pass mode (operation skip mode)
    - When the reference edge component value is negative

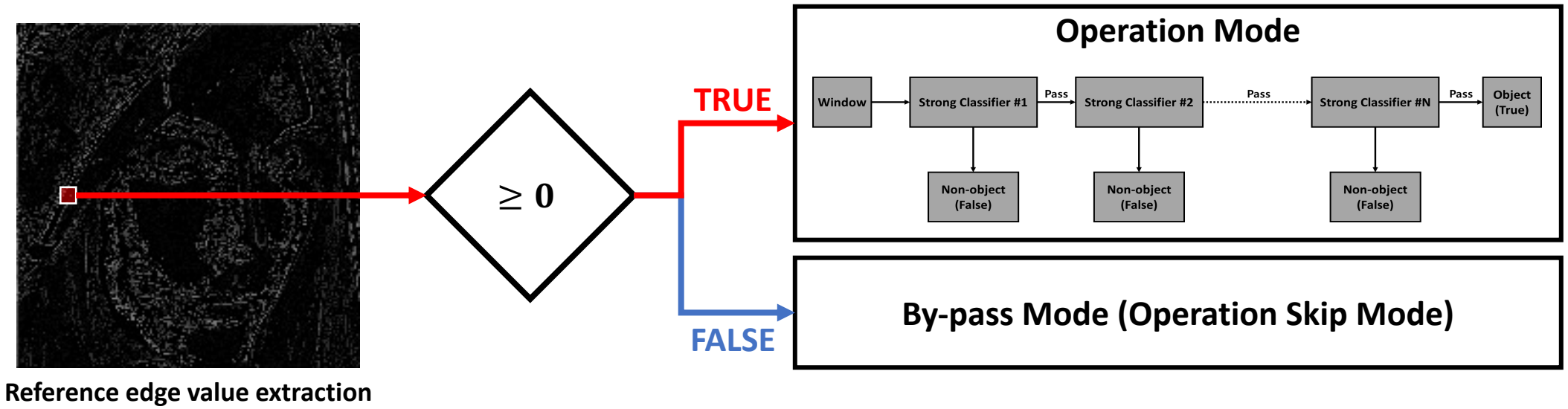


Fig 6. Proposed edge-based operation skip scheme process

4. Experimental Results

Experimental Results (1/4)

- **Environment**
  - ‘haarcascade_frontalface_alt.xml’
    - The number of strong classifiers: 22
    - Square window size: 20
  - Hyper-parameters
    - α : 0.5
    - β : 0.05
    - Intersection over union (IoU): 0.5
  - Test Frames
    - Lena
    - Solvay Conference 1927

Experimental Results (2/4)

- **Test Frame (Lena)**
  - Precision
    - Conventional: 12.5%
    - Hyun[5]: 12.5%
    - Choi[4]: 50%
    - Proposed: 100%
  - Recall
    - Conventional: 100%
    - Hyun[5]: 100%
    - Choi[4]: 100%
    - Proposed: 100%

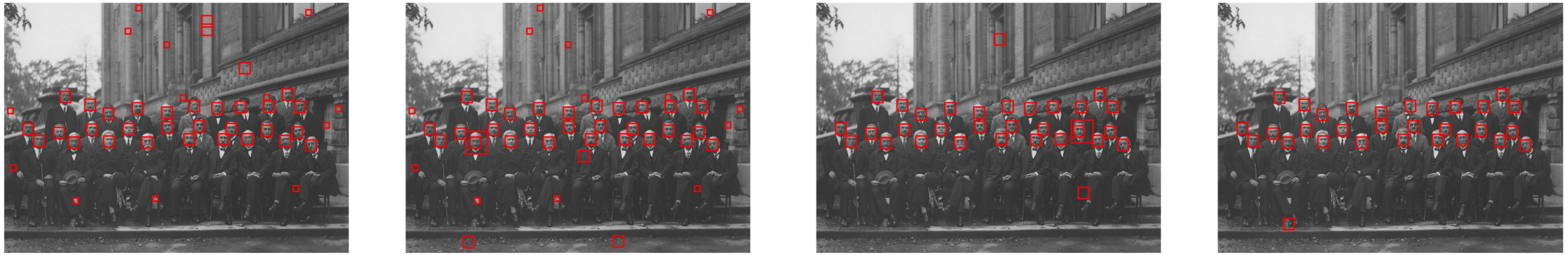
(a) (b) (c) (d)

Fig 7. Experimental results using the Lena test frame: **(a)** conventional, **(b)** Hyun, **(c)** Choi, and **(d)** Ours

[5] J. Hyun, et al., "Hardware architecture of a Haar Cascade Classifier based Face Detection System using a Skip Scheme," *in Proc. IEEE International Symposium on Circuits and Systems (ISCAS)*, 2021.

Experimental Results (3/4)

- **Test Frame (Solvay Conference 1927)**
  - Precision
    - Conventional: 64.44%
    - Hyun[5]: 64.44%
    - Choi[4]: 90.63%
    - Proposed: 96.67%
  - Recall
    - Conventional: 100%
    - Hyun[5]: 100%
    - Choi[4]: 100%
    - Proposed: 100%



(a) (b) (c) (d)

Fig 8. Experimental results using the Solvay Conference 1927 test frame: **(a)** conventional, **(b)** Hyun, **(c)** Choi, and **(d)** Ours

Experimental Results (4/4)

- **Iteration Number Comparison**
  - Lena
    - Reduction percentage (Minimum): 31.38% (vs. Choi[4])
    - Reduction percentage (Maximum): 84.73% (vs. Conventional)
  - Solvay conference 1927
    - Reduction percentage (Minimum): 36.63% (vs. Choi[4])
    - Reduction percentage (Maximum): 85.78% (vs. Conventional)

Table 1. The number of classification iteration when using conventional and proposed methods

Frame	Works			
	Traditional[2,3]	Hyun[5]	Choi[4]	Ours
Lena	15,502,079 (≅ **15.50M**)	15,480,375 (≅ **15.48M**)	3,450,202 (≅ **3.45M**)	2,367,604 (≅ **2.37M**)
Solvay Conference 1927	73,465,180 (≅ **73.47M**)	73,094,971 (≅ **73.09M**)	16,481,076 (≅ **16.48M**)	10,443,967 (≅ **10.44M**)

Conclusion

- Advantages
  - Proposed method can enhance processing speed by reducing the number of classification iterations
  - Proposed method exhibits better detection performance, consisting of precision and recall
- Limitation
  - Do not use public dataset
- Future Works
  - Conduct additional experiments using public datasets
  - Create dataset based on long-wave infrared (LWIR) thermal camera for satisfying CPD standard of Euro-NCAP

Q & A