

Research on Improving the Performance of Image based Web Structure Similarity: Combining SSIM and ORB algorithms

Seo-Hyuck Lee*, Jin-san Kim*, Jung-Hwan Kim*, Hanjin Lee**

*Researcher, Weven Corporation, Sejong, Korea

**Professor, School of Creative Convergence Education, Handong Global University, Pohang, Korea

[Abstract]

This study aims to establish a standard to accurately determine the similarity of the results when web pages are generated automatically using AI technology due to the explosive increase in demand for digital business. The YOLO, SSIM, Jaccard, and ORB techniques presented in previous studies related to the existing image similarity evaluation index generally focused on the partial and morphological similarity between the reference and the derived image. However, with the development of more complex and in-depth digital services based on generative AI, the need for comprehensive similarity analysis and determination methods that reflect the context and structure has emerged. Accordingly, this study proposed and verified a method to obtain 'Web Structural Similarity (WSS)' by combining the advantages of SSIM and ORB prior techniques. The research will serve various meaningful implications.

▶ **Key words:** YOLO, Image Similarity, Jaccard, SSIM, ORB, Structural Discrimination

[요 약]

본 연구는 폭증하는 디지털 비즈니스 수요 증가에 따라 AI 기술 등 자동화된 방식으로 웹 페이지가 생성됐을 때, 결과물의 유사도를 정확하게 판별하고자 기준을 수립하고자 한다. 기존의 이미지 유사도 평가지표와 관련된 선행연구에서 제시한 YOLO, FID, Jaccard, SSIM, ORB 기법들은 일반적으로 기준안과 파생된 이미지의 부분적, 형태적 유사도에 집중되어 있었다. 그러나 생성형 AI 기반의 더 복잡하고 심화된 디지털 서비스들의 발전에 따라 맥락과 구조를 반영한 종합적 유사도 분석, 판별방안이 필요함이 대두되었다. 이에 따라 본 연구에서는 SSIM과 ORB 선행기법들의 장점을 조합하여 '웹 구조적 유사도(WSS)'를 구하는 방식을 구상하여 제안 및 검증하였다. 연구결과로 개발된 알고리즘을 통한 생성 이미지 비교평가 시 유의미한 성능 개선에 시사점을 줄 것이다.

▶ **주제어:** YOLO, 이미지유사도, Jaccard, SSIM, ORB, 구조적판별

- First Author: Seo-Hyuck Lee, Corresponding Author: Hanjin Lee
- *Seo-Hyuck Lee (Ish@weven.kr), Weven Corporation
- *Jin-san Kim (kjs@weven.kr), Weven Corporation
- *Jung-Hwan Kim (kjhbond@gmail.com), Weven Corporation
- **Hanjin Lee (cus@handong.edu), School of Creative Convergence Education, Handong Global University
- Received: 2024. 09. 02, Revised: 2024. 10. 14, Accepted: 2024. 10. 30.

I. Introduction

최근 생성형 AI 기술의 급격한 발전은 웹 생성 및 콘텐츠 자동화의 새로운 가능성을 열어주고 있다[1]. 이러한 기술은 사용자 경험(UX)의 향상뿐만 아니라, 콘텐츠 생성 시간을 단축하고, 개인화된 웹 환경을 제공하는 데 중요한 역할을 한다. 특히, AI를 활용한 웹 디자인 및 콘텐츠 생성은 효율성을 극대화하며, 다양한 사용자 요구에 빠르게 대응할 수 있는 장점을 지니고 있다[2].

그러나 맞춤형 웹사이트 개설에 대한 수요가 지속적으로 증가함에도 불구하고, 전통적인 SI(System Integration)와 웹 에이전시들은 이 수요를 충족시키는 데 한계를 보이고 있다[1]. 프로젝트 기간의 지연과 높은 비용은 비효율성을 초래하며, 이러한 문제는 더욱 심화하고 있다. 이에 따라 생성형 AI 기술을 활용한 자동화된 웹 생성 솔루션이 주목받고 있지만, 현재의 기술은 그 품질과 신뢰성을 평가하는 데 있어 한계가 있다[2].

이와 같은 상황을 해결하고자 본 연구에서 제시하고자 하는 것은 입력한 웹 프레임 이미지와 생성된 웹 프레임 이미지 간의 유사도를 구하는 것이다. 현재도 두 이미지 간의 유사도를 구하는 Jaccard, SSIM, ORB 등의 이미지 유사도 평가기법들이 있지만, 주로 고정된 기준 이미지와 파생된 이미지 간의 부분적 비교에 초점을 맞추고 있어서 다양한 생성형 AI 결과물의 구조적 유사도를 평가하는 데는 한계가 있다. 이러한 한계는 특히 쇼핑몰 및 웹 페이지와 같은 구조적으로 복잡한 창작물에 대한 평가에서 특히 두드러진다[4].

예를 들어, 쇼핑몰 웹사이트의 경우, 다양한 상품 이미지와 텍스트 블록이 복잡하게 배치되어 있다. 이러한 웹 페이지에서 AI가 분석을 통해 만든 웹 프레임이 원본과 얼마나 유사한지를 평가하는 것은 매우 어렵다. 현재의 유사도 평가 방법들은 이러한 복잡한 구조를 충분히 반영하지 못해, 생성된 결과물이 실제로 사용자의 기대에 부합하는지 여부를 정확히 판단하기 어렵기 때문이다. 이는 쇼핑몰, 나아가 금융, 헬스케어 관련 사이트에서 고객의 신뢰도와 직결되기 때문에, 보다 정교한 평가방법이 필요하다.

한편 이미지 유사도를 측정하는 대표적인 방법으로는 평균 제곱 오차(MSE, Mean Squared Error), 최대 신호 대 잡음비(PSNR, Peak Signal-to-Noise Ratio), 구조적 유사도(SSIM, Structural Similarity Index Measure), FID(Frechet Inception Distance) 등이 있다[5,6].

각각의 방법은 이미지 분석의 특정 측면을 강조하는데, 장단점에 대한 선행연구 결과가 뚜렷하다. 이에 본 연구에서는 이러한 방법들을 종합적으로 결합하여 더 나은 유사도 측정방법을 개발하고자 한다.

나아가 본 연구에서는 생성형 AI를 활용한 웹 생성물의 구조적 유사도를 평가하기 위한 새로운 지표를 제안한다. 기존의 유사도 평가기법들이 고정된 기준 이미지와 파생된 이미지 간의 비교에 집중되는 반면, 본 연구에서는 다양한 생성형 AI 결과물 간의 구조적 유사도를 더욱 정확하게 평가할 수 있는 방법론을 개발한다. 이를 통해 웹 페이지와 같은 복잡한 구조의 이미지들에 대해 인지적으로 유사한 구조를 갖는지를 평가하는 데 중점을 둔다. 본 연구는 웹 생성 AI 기술의 평가지표로 활용될 수 있을 뿐만 아니라, 웹 디자인 및 콘텐츠 생성의 자동화 수준을 한 단계 높이는 데 기여할 것으로 기대된다.

II. Preliminaries

1. Related works

1.1 YOLO (You Only Look Once)

기존에 사용하던 웹페이지의 분류체계에 의한 비교는 정량적인 수치로 비교할 수 없어, 사용자(고객)의 요구사항에 얼마나 적합한 생성물이 탄생했는지 비교하기 어려웠다[3]. 또한 웹을 생성하는 퍼블리셔와 디자이너에 의해 얼마든지 자유롭고 변화무쌍한 형태가 제작될 수 있기에 사용자가 원하는 화면에 얼마나 유사하게 만들어 주는지에 대한 평가가 필요했다[2].

이에 가장 많이 쓰이는 알고리즘인 YOLO 객체탐지 모델을 적용해볼 수 있다[3]. 해당 방법론은 탐지된 객체의 영역에 대한 Confidence(일종의 신뢰도) 값을 기반으로 두 웹 페이지 간의 유사도를 판단한다(Fig 1). YOLOv8은 최신 객체탐지 모델로서, 실시간 탐지 성능을 극대화한 아키텍처를 제공한다. 이러한 특징 덕분에 다양한 웹 페이지 요소들(예: 이미지, 텍스트 블록 등)을 신속하고 정확하게 탐지할 수 있다[1]. 이에 따라, 탐지된 각 영역의 Confidence 값을 활용하여, 본 연구에서는 선행연구에서 구축한 분류체계[2]에 기반한 유사도 측정 방식을 통해 웹 페이지 전체의 구조적 유사성을 검토해보았다.

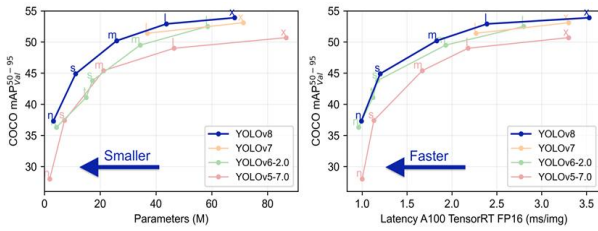


Fig. 1. Yolov8 Parameters(M), mAP Board [3]

그러나 이러한 접근 방식은 웹 페이지의 다양성과 디자인, 퍼블리셔의 개별적 스타일에 따른 변동성으로 인해 본질적인 한계에 직면하게 되었다. 웹 페이지 디자인은 기본적인 레이아웃에서부터 색상, 타이포그래피, 이미지 배치에 이르기까지 매우 다양한 형태를 취하며, 이러한 변화는 디자이너와 퍼블리셔의 창의적인 판단에 의해 크게 좌우된다[1]. 결과적으로, 동일한 콘텐츠나 기능을 담고 있는 웹 페이지일지라도 그 시각적 표현과 구조적 구성은 크게 달라질 수 있다.

이에 따라 YOLOv8 기반 탐지 방법은 웹 페이지 간 유사도를 평가하는 데 있어 일관적이고 신뢰할 수 있는 결과를 도출하는 데 어려움을 겪었다. 예를 들어, 탐지된 객체의 배치, 크기, 스타일이 다를 수 있음에도 불구하고, 기능적으로는 동일한 두 웹 페이지가 존재할 수 있는데, 이러한 YOLOv8의 Confidence(신뢰도) 값만을 이용해서는 이들 간의 유사성을 정확하게 평가하는 것이 어려웠다. 더 나아가, 웹 페이지의 비정형적인 디자인 요소들은 탐지 결과에 큰 영향을 미치며, 유사도 측정의 일관성을 저해하는 문제를 야기하였다[2,3].

결국, YOLOv8 모델을 활용한 유사도 측정 방법론은 구조적으로 복잡하고 변동성이 큰 웹페이지와 같은 생성물에 대해 정확한 비교분석을 제공하기 어렵다는 평가다. 이에 따라, 이러한 한계를 극복하고 더 정밀한 유사도 평가를 가능하게 할 새로운 접근 방법의 필요성이 대두되었다.

1.2 OpenCV

기본적으로 YOLOv8 모델은 입력 이미지 내의 주요 객체 및 요소들을 탐지한다. 이에 YOLOv8은 빠르고 정확한 객체탐지 능력을 갖추고 있으며, 웹페이지에서 텍스트, 이미지 영역 등과 같은 주요 요소들을 신뢰성 있게 감지할 수 있다. 감지된 각 영역은 Bounding box로 표시되며, 이를 기반으로 이미지의 구조를 이해할 수 있다[3]. 탐지된 Bounding box 영역을 기준으로, OpenCV를 사용하여 다음과 같은 전처리 작업을 수행한다[4].

1) 배경 색상 통일: 탐지된 영역 이외의 부분을 단일 색상(예: 회색)으로 통일한다[7]. 이를 통해, 탐지된 영역을

강조하고 비교 작업을 용이하게 한다[6].

2) 영역 크기 조정 및 정렬: 탐지된 각 영역의 크기를 동일하게 조정하고, 이를 특정 레이아웃(예: 좌상단 기준 정렬)으로 배치한다[7,8]. 이렇게 하면, 두 이미지 간의 구조적 비교가 더욱 명확해진다.

3) 탐지된 영역의 색상 변경: YOLOv8 모델로 탐지된 객체 영역은 각기 다른 색상으로 표시된다[3]. 예컨대 특정 객체가 포함된 영역을 특정 색상으로 채우고, 이 색상 코드를 이용해 다른 이미지와의 비교를 수행한다[6].

이러한 방법은 생성형 AI 기술로 제작된 웹페이지의 구조적 유사성을 평가하는데 중요한 역할을 한다[8]. OpenCV를 통해 일관된 이미지 전처리 과정(Fig 2)을 적용하고, YOLOv8과 같은 객체탐지 모델을 함께 활용함으로써, 불필요한 요소들이 유사도 평가에 유입되는 것을 막을 수 있다[4]. 이를 통해 웹페이지의 레이아웃, 객체 배치 등 구조적 요소들이 일관되게 평가될 수 있으며, 생성된 결과물의 품질에 대한 신뢰성을 높일 수 있을 것이다[6].



Fig. 2. Preprocessing Website Data by OpenCV

1.3 Perceptual Metric

대표적으로 FID(Frechet Inception Distance)와 IS(Inception Score) 등의 지표는 생성된 이미지의 시각적 품질과 사실감을 평가하는 데 주로 사용된다[5,6]. 이들은 이미지의 전반적인 품질을 수치화하여 생성된 이미지가 실제 이미지와 얼마나 유사한지를 측정한다. 일례로, FID는 생성된 이미지와 실제 데이터 간의 분포 차이를 측정하여, 이미지의 충실도(Fidelity)를 판단한다. 이러한 지표들은 고품질 이미지의 사실감을 평가하는 데 유리하지만, 웹 페이지의 구조적 요소를 분석하는 데는 한계가 있다.

웹 페이지는 이미지뿐 아니라 텍스트, 레이아웃, 메뉴 구조, 인터랙티브 요소 등이 조합된 복잡한 구조물입니다. 웹 페이지의 모서리 정렬, 이미지와 텍스트 간의 배치 간격, 레이아웃의 일관성 등은 사용자경험(UX)에 큰 영향을

미치는 중요한 요소들이지만, FID와 IS는 이러한 세부적인 구조적 특징을 반영하지 못한다[5].

또한, FID는 비교 대상의 해상도가 다를 경우 평가가 어려운 문제가 있다. 웹페이지는 다양한 해상도와 반응형 요소에 따라 디자인이 조정되기 때문에, 정확한 결과를 제공하지 못할 가능성이 크다. 이러한 한계는 구조적 유사성을 평가하는 데 있어 큰 제약이 된다[6].

나아가 FID와 IS는 훈련된 데이터 세트 내에서 유사한 이미지를 생성하는 특성을 기반으로 하기 때문에, 입력값이 다양하게 변화할 수 있는 웹 페이지 생성 AI의 평가에는 적합하지 않다. 이들 지표는 고정된 데이터셋을 전제로 하여, 웹 페이지처럼 동적으로 변하는 콘텐츠의 유사성을 평가하기에는 한계가 있다[5,6].

1.4 JSI (Jaccard Similarity Index)

그러나 구조적 유사성 평가의 정확성을 더욱 높이기 위해서는 다양한 측면에서 접근할 필요가 있다. 실제로, 페이지의 구조뿐만 아니라 요소 간의 위치적 유사성도 중요한 평가 요소가 될 수 있다. 이에 따라, 본 연구에서는 추가로 Jaccard 유사도 평가(Jaccard Similarity Index, 이하 JSI) 방법을 검토하였다[9]. 이는 두 집합 간의 유사성을 측정하는 통계적 방법으로, 주로 집합이론에서 사용된다. 즉, 두 집합 간의 공통요소 비율을 기반으로 유사성을 계산(Fig 3)하며, 0과 1 사이의 값을 가진다[9]. 1에 가까울수록 두 집합이 매우 유사하다는 것을 의미한다.

$$Jaccard \text{ 유사도} = \frac{|A \cap B|}{|A \cup B|}$$

Fig. 3. Jaccard Similarity Calculation formula

여기서 A와 B는 비교하고자 하는 두 집합이고, 두 집합 간의 교집합에 속하는 요소의 개수, 두 집합 간의 합집합에 속하는 요소의 개수를 나타낸다[7]. Jaccard 유사도는 텍스트 유사도, 이미지 분할 평가, 객체탐지 결과평가 등 다양한 분야에서 활용된다.

이에 본 연구자들은 생성형 AI로 생성된 웹 페이지 이미지의 유사도를 평가하기 위해 Jaccard 유사도를 활용하고자 검토하였다. YOLOv8 모델을 사용하여 웹 페이지 내의 다양한 객체(예: 버튼, 이미지, 텍스트 등)를 탐지하고, 이 탐지된 영역들을 각각의 집합으로 간주한 뒤, 본 유사도를 통해 웹 페이지 간의 구조적 유사성을 평가해보았다.

예컨대, 두 웹 페이지 이미지에서 YOLOv8이 탐지한 객체들의 좌표를 사용해 각각의 탐지결과를 집합 A와 B로

나타낼 수 있을 것이다. 그 후, 두 집합 간의 Jaccard 유사도를 계산하여 두 웹 페이지가 얼마나 유사한지를 수치상으로 평가($t1=0.63$)하였다.

그러나 이 결과값에 대해 선행연구에 따르면 해석에 주의를 요한다[7,9]. 수치가 무조건 높다고 유사도가 높고, 반대로 낮은 수치는 유사하지 않다고 해석하기에는 한계가 존재한다. 왜냐하면 웹 페이지의 디자인 요소는 디자이너의 개별적 판단에 따라 크게 달라지기 때문에, 같은 기능이나 내용을 가진 웹 페이지라도 그 구조적 배치나 시각적 표현이 매우 다를 수 있다. 이런 경우, 탐지된 객체 간의 단순한 교집합과 합집합만을 고려하는 Jaccard 유사도는 웹 페이지 간의 본질적인 웹페이지의 구조적인 유사성을 제대로 반영하지 못했다.

또한, 웹 페이지는 매우 다양한 요소들로 구성되며, 이 요소들은 크기, 위치, 스타일이 전혀 다를 수 있다. Jaccard 유사도는 이러한 복잡성을 충분히 반영하지 못해, 결과적으로 두 웹 페이지 간의 유사성을 정확하게 평가하는 데 한계가 있었다.

1.5 SSIM (Structure Similarity Index Measure)

다음으로 검토한 선행연구 기법은 SSIM(구조적 유사도 지표 측정법)으로, 이미지 간의 유사성 및 품질평가에 사용된다[8]. SSIM은 단순한 픽셀 기반 비교보다는 인간의 시각 시스템이 이미지를 인지하는 방식에 근접한 방법으로 유사성을 평가한다. 이 지표는 두 이미지의 밝기, 대비, 구조를 비교하여 유사성을 계산하며, 0에서 1 사이의 값을 가진다[7]. SSIM의 값이 1에 가까울수록 두 이미지가 매우 유사하다는 것을 의미한다. SSIM은 다음과 같은 세 가지 요소를 결합하여 유사도를 계산한다[8].

1) 밝기(Luminance, $l(x, y)$): 두 이미지의 평균 밝기(휘도)를 비교하며, SSIM 계산 시 이미지의 픽셀값을 이용한다. 즉, 픽셀값이 클수록 밝음을 이용하며, 별도 빛의 밝기 성분을 추출해서 사용하지는 않는다. 회색조(gray-scale) 이미지에서는 각 픽셀의 값을 의미하며, RGB 이미지에서는 R, G, B 각 채널별 픽셀값을 의미한다.

2) 대비(Contrast, $c(x, y)$): 두 이미지의 빛의 밝기가 바뀌는 정도(대조)를 비교한다. 픽셀 간의 값이 얼마나 차이가 나는지 정량화할 수 있으므로 표준 편차를 사용한다.

3) 구조(Structure, $s(x, y)$): 두 이미지 픽셀값의 구조적인 차이점을 기반으로 유사성을 비교하는 방식이다. 이

를 구하기 위하여 '밝기(luminance)'를 평균, '대비(contrast)'를 표준 편차로 이용하여 정규화(Normalized)된 픽셀값의 분포에서 측정치를 재정의한다[4].

종합적으로 SSIM은 <Fig 4>의 수식으로 계산된다.

$$SSIM(x, y) = [l(x, y)]^\alpha \cdot [c(x, y)]^\beta \cdot [s(x, y)]^\gamma$$

Fig. 4. SSIM Calculation Formula [4]

아울러 각각의 요소는 다음 수식(Fig 5)으로 계산한다.

$$\begin{aligned} \text{밝기: } l(x, y) &= \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1} \\ \text{대비: } c(x, y) &= \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2} \\ \text{구조: } s(x, y) &= \frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3} \end{aligned}$$

μ_x, μ_y : 각각두 이미지의 평균 밝기
 σ_x, σ_y : 각각두 이미지의 표준 편차(대비를 나타냄)
 σ_{xy} : 두 이미지 간의 공분산(구조적 유사성을 나타냄)
 C_1, C_2, C_3 : 안정화상수로, 아주 작은 값으로 설정하여 분모가 0이 되는 것을 방지

Fig. 5. Specific Calculate Formula of SSIM [4,8]

웹페이지의 일부분을 SSIM을 통해 생성된 결과와 비교하는 방법도 이리하다(Fig 6). 생성된 웹페이지, 비교 대상 웹페이지를 이미지화하고 SSIM 유사도 평가 프로세스에 입력해 결과를 확인한다[8]. 값이 1에 가까울수록 유사한 것으로 판단한다.

하지만 밝기와 대비를 비교하는 부분(Table 1과 Fig 7)에서 채도, 명도, 조도에 영향을 받기에 구조적인 유사도만을 평가하기는 어렵다는 한계가 있었다[6,8].

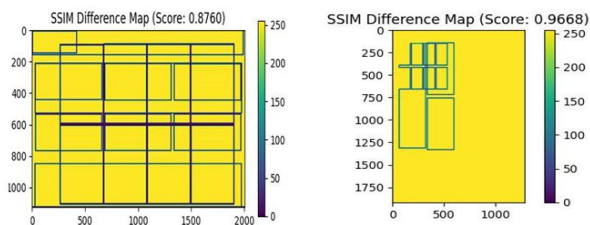


Fig. 6. SSIM Algorithm Structural Area Calculation Results

Table 1. Results comparison analysis by color

Background \ Image	(10, 10, 10)	(100, 100, 100)	(180, 180, 180)	(245, 245, 245)
B_(0,0,0)	0.6989	0.6197	0.6151	0.6136
W_(255, 255, 255)	0.6454	0.7943	0.8665	0.9495

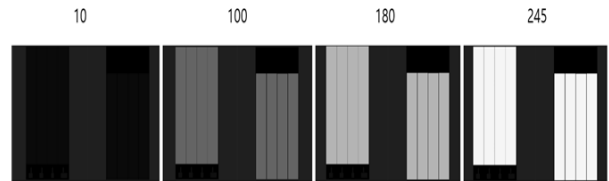


Fig. 7. Limitation of SSIM (Example)

1.6 ORB (Oriented FAST and Rotated BRIEF)

이러한 한계를 극복하기 위해, SSIM에서 나아가 보다 구조적 유사성에 초점을 맞춘 ORB 알고리즘 사용을 시도하였다. ORB는 (Oriented FAST and Rotated BRIEF) 컴퓨터 비전 분야에서 널리 사용되는 특징점 검출 및 Descriptor 생성 알고리즘이다[10]. ORB의 특징은 빠르고, 효율이면서도 강력한 성능을 제공한다. 주로 객체 인식, 이미지 정합, 추적 등의 작업에서 사용한다.

왜냐하면 이는 특징점 기반 접근법을 통해 이미지 간의 유사성을 더욱 정밀하게 분석할 수 있기 때문이다. 이미지 내의 특정 위치에서 특징점을 추출하고, 이들 간의 위치적 일치성을 기반으로 유사성을 평가하는데 강점을 가진다. SSIM과 달리, ORB는 밝기와 대비 요소를 배제하고 구조적 특징점에 중점을 둔 평가를 가능하게 한다[10].

ORB는 크게 두 가지 주요 알고리즘을 결합하여 동작한다. 먼저 FAST(Features from Accelerated Segment Test)는 ORB에서 사용되는 특징점(코너 등)을 검출하는 알고리즘이다[10]. FAST는 매우 빠르게 특징점을 찾을 수 있지만, 이 특징점들이 회전이나 조명 변화에 민감할 수 있다. ORB는 이 단점을 보완하기 위해 FAST로 검출된 특징점에 대해 방향 정보를 추가로 계산하여 회전 불변성을 갖도록 한다.

한편 BRIEF(Binary Robust Independent Elementary Features)는 특징점 주변의 패치를 이진화하여 설명하는 기술자(Descriptor)를 생성하는 알고리즘이다[10]. BRIEF는 계산이 매우 간단하고 빠르지만, 역시 회전 불변성을 갖추고 있지 않다. ORB는 BRIEF를 회전 불변하게 개선하여, 다양한 각도에서 촬영된 이미지에서도 동일한 특징점을 찾아낼 수 있도록 한다.

ORB 알고리즘의 동작 과정은 다음과 같다[10].

1) 특징점 검출: 이미지 내의 코너나 엣지와 같은 특징점을 FAST 알고리즘을 통해 검출한다. 이 특징점들은 이미지의 중요한 픽셀을 나타내며, 이미지의 주요 구조를 파악하는 데 중요한 역할을 한다.

2) 특징점 방향 계산: 검출된 특징점 주변의 화소 분포를 분석하여 각 특징점의 방향을 계산한다. 이를 통해 ORB 알고리즘은 회전 불변성을 가지게 되며, 이미지가 회전해도 일관된 특징점을 유지할 수 있다.

3) 기술자(Descriptor) 생성: BRIEF 알고리즘을 사용하여 특징점 주변의 패치를 이진화하고, 이를 기반으로 기술자를 생성한다. 이는 매우 짧고 계산이 효율적이며, 이미지의 주요 특징을 포착하여 시각화하는 데 사용된다.

4) 특징점 매칭: 다른 이미지에서 동일한 개체나 장면을 찾기 위해 생성된 기술자들을 비교하여 두 이미지 간의 유사한 특징점을 매칭한다. 이 과정은 일반적으로 해밍 거리(Hamming Distance)를 사용하여 수행된다.

이와 같은 ORB 알고리즘의 동작 과정을 통해 두 이미지 사이의 특징점을 추출하고, 이들이 얼마나 일치하는지를 분석한다. ORB는 특히 이미지 내의 중요한 구조적 요소인 코너(corner)를 효과적으로 검출하고, 이들 간의 유사성을 비교하는 데 강점을 가진다. 일례로 웹페이지의 레이아웃에서 중요한 역할을 하는 다양한 코너들을 비교할 수 있다(Fig 8, Fig 9). 이에 따라 구조적인 부분에서 각 요소가 적절하게 배치되었는지, 각 코너의 위치가 예상된 오프셋(offset)을 가졌는지를 판단할 수 있다[6,7].

이 과정에서 ORB는 코너 영역에 있는 특징점(Key Point)을 찾기 위해 해리스 코너 검출 알고리즘(Harris Corner Detection)을 사용한다. 해리스 코너 검출은 이미지에서 코너를 감지하는 고전적인 방법으로, 이미지의 경계를 찾기 위해 소벨(Sobel) 미분 연산을 활용한다. 소벨 미분은 이미지의 경계값을 계산하고, 이 경계값의 경사도 변화를 측정하여 코너를 검출하는 방식으로 동작한다[7]. 경사도가 수직, 수평, 대각선 방향으로 크게 변화하는 부분을 코너로 판단하여, 이미지 내의 구조적 특징을 포착하는 데 사용된다[9].

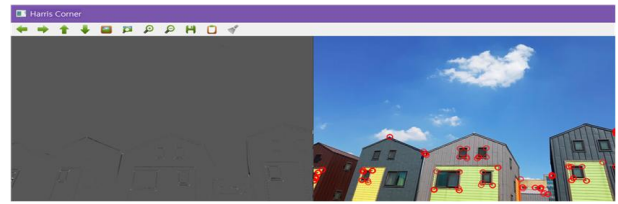


Fig. 8. Example of ORB

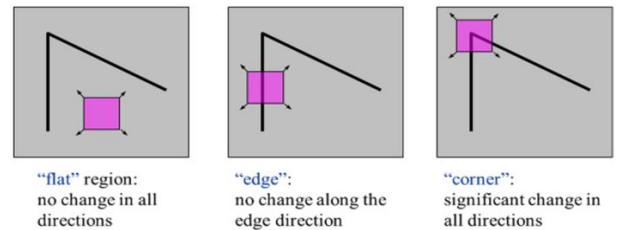


Fig. 9. How ORB finds Different Feature Points

따라서, 본 연구에서는 ORB 알고리즘을 사용하여 코너 영역의 특징점을 검출하고, 이들 간의 유사성을 분석함으로써, 웹페이지와 같은 구조적으로 복잡한 이미지 간의 유사성을 평가하고자 한다. 특히, 해리스 코너 검출과 시-토마시(Shi-Tomasi) 검출 알고리즘을 함께 고려하여, 다양한 조건에서의 이미지 유사성을 더욱 정확하게 평가할 방법을 제안한다[11]. 이를 통해, AI가 생성한 웹페이지가 원본과 얼마나 일치하는지를 평가하는 데 있어, 기존의 유사도 평가를 보완하고 더욱 정교한 분석을 가능하게 할 것이다.

III. The Proposed Scheme

앞서 밝힌 바와 같이 본 연구에서는 다양한 유사도 검증 기법들인 Jaccard, SSIM, ORB 등을 평가지표로 검토해보았다. 그러나 해당 기법들을 통해 얻은 결과값이 실제 인적 판단과 상당히 상이하여, 웹페이지 구조적 유사성 평가에 직접적으로 적용하는 데 어려움이 있었다. 이러한 한계를 극복하기 위해, 기존 알고리즘의 이론적 기반을 토대로 여러 기법들을 개선하고 결합하여 사용한다면 더 유의미한 결과를 도출할 수 있을 것으로 판단하였다. 이에 따라, ORB의 특징점 추출 방식과 SSIM의 구조적 유사성 판단 부분을 결합하여 활용하는 접근 방식을 제안한다. 이와 같은 결합된 접근법을 통해 웹페이지의 시각적 및 구조적 유사성을 보다 정교하게 평가할 수 있을 것으로 기대된다.

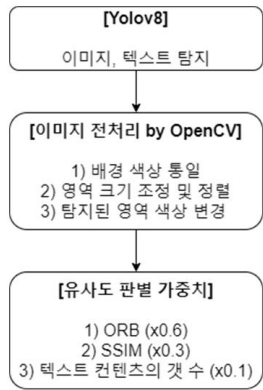


Fig. 10. Overall Workflow of Similarity Calculate

3.1 Detection Corner

먼저 영역을 검출하는 과정에서 특정 영역을 더 정밀하게 분석하기 위해, 우리는 이미지 내에서 코너의 특정 모양을 지정하여 분석하였다. 이러한 코너들은 웹페이지 레이아웃에서 중요한 구조적 요소로 작용하며, 다양한 형태를 가질 수 있다. 실제로 직각 코너(┌, ┐, └, ┘), 교차점(+), 그리고 T자형 모양(┌, ┐, └, ┘) 등 다양한 코너 형태가 있을 것이다. 이와 같은 코너들은 웹페이지의 구조적 일관성을 평가하는 데 중요한 역할을 하며, 각 코너의 위치와 형태를 비교하여 구조적인 오프셋(offset)이 발생했는지를 판단하는 데 사용된다[12].



Fig. 11. Criteria for being Recognized as a Corner

특히, 이들 코너 형태의 특징을 기반으로, 두 이미지 간의 구조적 유사성을 더욱 정밀하게 평가할 수 있다[11]. 이를 위해, 해당 코너들에 대한 정확한 위치와 구성을 분석하고, 이미지 내에서 동일한 위치에 유사한 형태의 코너가 존재하는지를 비교하였다. <Fig. 11>과 같이 ┌ 모양의 코너가 특정 위치에 존재한다면, 비교 이미지에서도 동일한 위치에 ┌ 모양의 코너가 존재하는지를 확인하는 것이다. 이러한 비교를 통해, 두 이미지 간의 구조적 일치도를 더욱 명확하게 평가할 수 있다[12].

이와 같은 코너 기반의 분석 방법은 웹페이지와 같은 구조적으로 복잡한 이미지에서 매우 유용하다[10]. 코너의 형태와 위치를 정확하게 비교함으로써, 단순히 픽셀 단위의 유사성을 넘어, 이미지의 전반적인 구조적 유사성을 평가할 수 있게 된다. 특히, 웹페이지 레이아웃에서 중요한

역할을 하는 이러한 코너들은 전체적인 디자인의 일관성을 유지하는 데 중요한 요소로 작용하므로, 이를 통해 보다 정밀한 구조적 유사성 평가가 가능해진다[9].

또한, 지정된 코너 모양을 활용한 분석은 웹페이지의 시각적 안정성을 평가하는 데에도 기여할 수 있다. 예컨대, 웹페이지 내의 특정 코너가 예상 위치에서 벗어나 있거나, 다른 형태로 변형되어 있으면, 이는 레이아웃의 불안정성을 나타낼 수 있으며, 이를 조기에 발견하여 수정할 수 있다. 이처럼, 다양한 코너 형태를 기반으로 한 분석은 단순한 유사성 평가를 넘어, 웹페이지의 구조적 완성도와 시각적 일관성을 높이는 데 중요한 역할을 한다[1].

이러한 코너 기반의 접근법은 기존의 단순한 픽셀 비교 방식에서 벗어나, 더 깊이 있는 구조적 분석을 가능하게 하며, 특히 웹페이지 디자인과 같은 복잡한 구조를 평가하는 데 있어 중요한 도구로 활용될 수 있다[2].

한편 구조적인 유사도를 평가할 때, 동일하거나 유사한 위치에 나타난 특징점 간의 매핑이 중요한 문제로 대두되었다. 동일한 위치에 동일한 특징을 매칭하지 않으면, 유사도 계산에서 혼란이 발생할 수 있다. 이에 따라, 이미지의 코너에 위치한 특징점이 다른 이미지에서 비슷한 위치에 해당하지 않을 경우 매칭하지 않도록 해야한다.

3.2 Developing the Model

이를 위해, 두 이미지의 좌표 기준으로 매칭된 특징점의 깊이 유클리디안 영역(변수)을 벗어나지 않도록 제한했다. <Fig 12>과 같이 유사한 위치에 있는 특징점들만을 매칭하는 커스터마이징 코드를 개발했다[13].

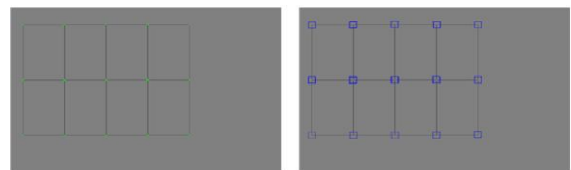


Fig. 12. Identifying Feature Points using (co)sine Values of Circular Patterns and Points using Rectangular Patches

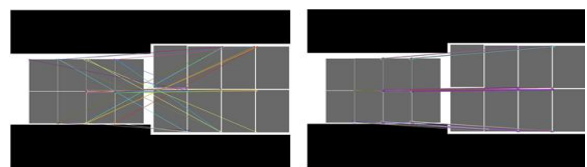


Fig. 13. Mapping only Feature Points in Similar Locations based on two Image Coordinate Values(Euclidean distance)

첫째, 원형 패턴의 Cosine, Sine 값을 사용하지 않고, 직사각형(Rectangle) 형태의 패치를 기준으로 이동하며

특징점을 확인한다. 둘째, 두 이미지의 좌표 값(유클리디언 거리)을 기반으로 유사한 위치에 있는 특징점만을 매칭한다[13]. 위 두 가지 조건을 모두 만족하는 경우에만, 양 이미지 간의 구조적인 특징점을 매핑한다(Fig 13).

다음으로 두 이미지에서 유사한 위치에 있는 특징점들만을 추출하고, 이를 연결하여 구조적인 유사도를 높이는 방식으로 커스터마이징된 매칭 알고리즘을 시각적으로 나타낸다[14]. 이 알고리즘은 매칭된 특징점이 유클리디언 거리 50~100 사이에 있는 경우만을 추출하여 구조적 유사도를 계산한다.

게다가, Descriptor 계산 방식에서 원형 패턴 대신 직사각형 패턴을 적용하여, Keypoints에서 사각형 영역에 적합한 Descriptor를 추출하는 방식을 사용하였다[13,14]. 이를 통해 두 이미지를 비교할 때 내부 요소들의 간격(Interval)을 조절하여 구조적 유사성을 강조하고, 시작점의 위치를 동일하게 설정하여 비교의 정확성을 높였다.

이 커스터마이징된 알고리즘은 유사한 위치에 있는 특징점들만을 고려함으로써, 구조적 유사도를 더욱 정확하게 평가할 수 있도록 설계되었다. 이로써, 이미지 간의 구조적 일관성을 더욱 신뢰성 있게 분석할 수 있게 되었다.

3.3 ORB and SSIM Combined Model

그러나 ORB를 단독으로 사용하였을 때, 이미지의 코너와 같은 특징점은 잘 분석되었지만, 이미지 내에서 중요한 역할을 하는 영역의 구조적 요소들을 충분히 반영하지 못하는 한계가 있었다[5,7]. 특히, 웹페이지와 같은 복잡한 이미지에서는 코너 외에도 다양한 요소가 유사성 평가에 영향을 미친다[8]. 이를 보완하기 위해, ORB와 SSIM을 결합하여 사용하기로 하였다.

즉, SSIM은 이미지의 밝기, 대비, 구조(겹치는 부분)를 함께 고려하여 평가할 수 있기 때문에, ORB로는 놓칠 수 있는 내부 영역의 유사성을 보완하는 데 유용하다[9]. 두 알고리즘을 결합함으로써, 웹페이지의 구조적 유사성을 평가하는 동시에, 내부의 채워진 영역과 같은 세부 요소까지도 함께 평가할 수 있었다. 이를 위해 2가지 지표의 비율과 가중치 설정을 위해 300개의 테스트 세트를 생성하여 측정하였다. 그 결과 이미지의 구조적 유사도를 구하는 본 실험의 특성 상, ORB와 SSIM의 가중치를 2:1로 잡는 경우에 사람의 인지수준에 가장 비슷한 수치를 나타내었다.

한편 테스트 과정에서 일정 수준 이상의 만족할만한 값이 도출되지 않는 점에 대해 탐구한 결과, 이미지 내 일부 텍스트 영역으로 인한 ORB 값의 변화가 한계를 보였다. 구조적 유사도를 봐야하는 본 연구의 특성 상 ORB의 가중

치가 높을 수 밖에 없는데, 텍스트의 영향을 줄이기 위해 가중치를 낮추기에는 전체 결과에 지대한 영향을 주기 때문에 해당 부분에 대한 추가적인 조치가 필요했다[10].

페이지의 구조적 유사성을 평가하는 데 있어 텍스트 영역은 중요한 요소로 작용한다[3]. 웹페이지에서 텍스트는 단순한 정보 전달의 역할을 넘어, 페이지의 전반적인 디자인과 사용자 경험(UX)에 큰 영향을 미친다. 하지만 이는 UX 측면에서 큰 영향을 주는 것이고, 유사도 측면에서는 영향이 낮다고 판단했다. 이에 텍스트 영역은 별도로 각 이미지 내 포함된 텍스트 영역의 갯수를 비교하는 가중치(0.1)를 포함하여 분석 진행하고자 했다[5,8].

이와 같은 과정에 따라서 “ORB : SSIM : 텍스트 갯수 = 0.6 : 0.3 : 0.1”의 가중치를 부여하였다. 이는 웹페이지의 구조적 유사성을 평가하는데 있어 텍스트와 그 외의 시각적 요소들을 균형있게 고려하도록 설계된 평가 방법론이다. 전체 수식은 <Fig. 14>와 같으며 세 가지 주요 요소를 결합하여 전체 유사성을 계산한다. 각 요소는 웹페이지의 요소별 중요한 측면을 담아내며, 수식은 이 요소들을 가중치와 함께 조합하여 최종적인 유사성 값을 도출한다.

$$\left(\frac{\text{Min}(\text{ElementsDetect}(\text{Counts}))}{\text{Max}(\text{Elements.ALL}(\text{Counts}))} \times \text{Counts} \right) + \left(\frac{\text{PositionMatchingCount}}{\text{compareImageElements} \cap \text{MatchingImageElements}} \times \text{Positions} \right) + \left(\frac{\text{Min}(\text{Text}(\text{Counts}))}{\text{Max}(\text{Text}(\text{Counts}))} \times \text{Texts} \right)$$

Fig. 14. Web Structure Similarity(WSS) Calculate Formula

수식 내 ElementsDetectCounts는 감지된 요소의 수를 의미하며 ElementsAllCounts는 모든 요소의 수를 의미한다. 이 비율은 감지된 요소 수를 전체 요소 수로 나눈 값으로, 요소 감지의 포괄성을 평가한다[6,9]. 아울러 Counts(ω)는 가중치를 의미하며, 이 비율에 곱하여 최종 점수를 산출한다. PositionMatchingCount는 위치 일치치의 횟수를 지칭하며, $\text{compareImageElements} \cap \text{MatchingImageElements}$ 는 입력 이미지와 일치하는 이미지 요소들의 교집합을 의미한다. 이 비율은 두 이미지 간의 위치 일치성을 평가한다. 한편 Position(ω): 이 비율에 곱해지는 가중치이다. TextCounts: 텍스트 요소의 수를 의미하고, 이 비율은 텍스트 요소의 일관성을 평가하며, text(ω): 텍스트 요소의 유사성에 부여된 가중치다.

이 세 요소의 가중 합계를 통해 웹페이지의 전체 구조적 유사성을 0에서 1 사이의 값으로 계산할 수 있다. 즉, 이

값이 1에 가까울수록 두 웹페이지의 구조가 유사하다고 평가할 수 있다[4,7]. 이를 통해 AI가 생성한 웹페이지와 원본 웹페이지 간의 유사성을 평가하는 데 있어, 보다 정교한 결과를 도출할 수 있을 것이다[5]. <Fig 15>에 제시된 케이스들을 토대로 적용해본 결과, <Table 2>에서 보는 바와 같이 조합된 모델 수치가 가장 높았다. 본 결과는 웹 페이지 간의 구조적 유사성을 객관적이고 신뢰성 있게 평가할 수 있는 기준을 제공한다.

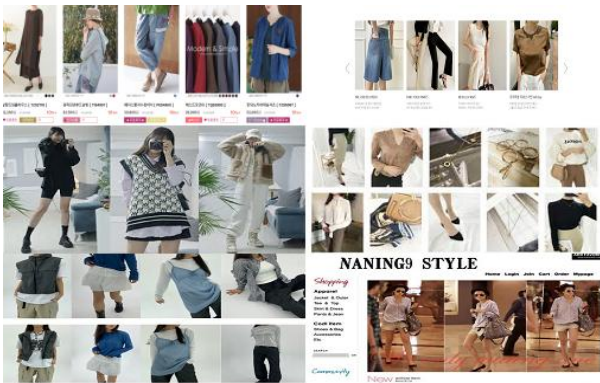


Fig. 15. Case Data 1, 2, 3 Illustrative Example

Table 2. Evaluation Values using Real Data

Property	Kp1	Kp2	Mp	ORB.S	SSIM.S	Comb.S
Case.1	90	63	38	0.60	0.65	0.70
Case.2	90	71	27	0.38	0.44	0.65
Case.3	90	33	10	0.31	0.35	0.59

Kp : Key Point (* ORB 기법으로 감지된 포인트)
 Mp : Matching Point (* Kp1과 Kp2 사이에 매칭된 포인트)
 ORB.S : ORB Similarity
 SSIM.S : SSIM Similarity
 Comb.S : Combine Similarity

IV. Conclusions

본 연구는 AI가 자동으로 생성한 웹 페이지와 기준 페이지 간의 구조적 유사성을 평가하기 위해 다양한 기법들을 탐색하였다. YOLO, FID, Jaccard, SSIM, ORB 기법들은 이미지의 부분적, 형태적 유사도에 집중되어 있었다. 그러나 AI 기반의 더 복잡하고 심화된 디지털 서비스들의 발전에 따라 맥락과 구조를 반영한 종합적 유사도 분석법이 필요했다. 모바일과 웹의 사용량 증가 및 디지털 기술 고도화로 AI 생성 웹 페이지가 증가하는데, 이 품질을 객관적으로 평가함으로써, 기획과 디자인 관점에서 사용자 경험이 얼마나 잘 반영되었는지를 판단할 필요가 있다. 이를 통해, 웹 디자인 자동화의 수준을 높이고, 다양한 산업 분야에서 생성형 AI의 활용 가능성을 증대시키고자 한다[15].

이에 본 연구결과, 기존의 ORB와 SSIM 기법을 가중치로 결합한 방법론이 더 정확한 구조적 유사도 평가를 제공할 수 있음을 확인했다. ORB는 구조적 특징점을 효과적으로 검출하고, SSIM은 이미지의 시각적 일관성을 평가하는데 강점을 지니고 있어, 두 알고리즘의 결합을 통해 보다 포괄적인 평가가 가능해졌다. 그러나, 각 기법의 결합 사용에 대한 구체적인 근거를 제시하는 데 있어 연구 초기의 어려움이 있었으며, 웹 페이지 디자인의 복잡성과 주관적인 요소를 평가하는 데 한계가 있었다.

연구 설계에 있어 웹 페이지 디자인의 주관적 성격과 다양한 디자이너의 변동성을 반영하는 데 어려움이 있었다. 또한, 상대적으로 미지에 남아있는 영역들이 많아 테스트 세트를 구성하여 실험적 척도를 사용해 결과를 제시하는 한계가 존재했다. 아울러 가중치 설정에 있어 구체적인 결합 근거를 완벽히 설명하기 난관도 여전히 남아있다.

향후 연구에서는 '구조적 유사도'의 측정을 더욱 정밀하게 하기 위해, 기존의 구조적 척도들을 결합하거나 텍스트 영역을 함께 고려함으로써 더 높은 신뢰도의 평가 방법을 개발할 수 있을 것이다. 또한, 생성형 AI를 이용해 생성된 이미지를 반응형 웹 페이지로 변환하고, 이를 평가할 수 있는 새로운 지표들을 도입함으로써, AI 기반 웹 페이지 생성의 자동화 수준을 한 단계 높일 수 있을 것이다.

결과적으로, 본 연구에서 제안한 방법론은 AI 생성 웹 페이지의 품질평가에 있어 더 정확하고 신뢰성 있는 결과를 제공하는데 기여할 수 있을 것이다. 이는 향후 AI 기술의 발전과 웹 디자인의 자동화를 선도하는 중요한 기준으로 학술적, 실무적 시사점을 제시한다.

REFERENCES

- [1] Kim, J.-H., Ko, Y., Choi, J., and Lee, H. (2024). "Research on the Design of a Deep Learning-Based Automatic Web Page Generation System", *Journal of the Korea Society of Computer and Information*, 29(2), 21-30. DOI : <https://doi.org/10.9708/JKS.CI.2024.29.02.021>
- [2] Kim, J.H., Cho, J.W., Kim, J.S., and Lee, H.J. (2024). "Research on Training and Implementation of Deep Learning Models for Web Page Analysis", *The Journal of the Convergence on Culture Technology*, 10(2), 517-524. DOI : <https://doi.org/10.17703/JCC.T.2024.10.2.517>
- [3] Juan, T. and Diana, C-E. (2023). "A Comprehensive Review of YOLO Architectures in Computer Vision: From YOLOv1 to YOLOv8 and YOLO-NAS", DOI : <https://doi.org/10.3390/make5040083>

- [4] Lee, J-Y., and Kim, Y-J. (2016). "Research of Improved Perceptual Image Quality Assessment based on Structural Similarity", DOI : <https://dspace.ajou.ac.kr/handle/2018.oak/12128>
- [5] Jung, J-M., Yang, H-J., Kim, S-H., Lee, G-S., and Kim, S-H. (2011) "Wine Label Recognition System using Image Similarity", 11(5), 125-137, DOI : <https://doi.org/10.5392/jkca.2011.11.5.125>
- [6] Wang, Z., Bovik, A.C., Sheikh, H.R., and Simoncelli, E.P.(2004). "Image Quality Assessment: From Error Visibility to Structural Similarity", 13(4), 600-612, DOI : <https://doi.org/10.1109/TIP.2003.819861>
- [7] Chen, G-H., Yang, C-L., Po, L. and Xie, S-L. (2006). "Edge-Based Structural Similarity for Image Quality Assessment", Conference on Acoustics, Speech, and Signal Processing, DOI : <https://doi.org/10.1109/ICASSP.2006.1660497>.
- [8] Renieblas, G.P., Nogués, A.T., González, A.M., Gómez-L, N., and Castillo, E.G. (2017). "Structural similarity index family for image quality assessment in radiological images", DOI : <https://doi.org/10.1117/1.JMI.4.3.035501>
- [9] Bag, S., Kumar, S. K., and Tiwari, M. K. (2019). "An Efficient Recommendation Generation using Relevant Jaccard Similarity", Information Sciences, 483, 53-64, DOI : <https://doi.org/10.1016/j.ins.2019.01.023>
- [10] Karami, E., Prasad, S., and Shehata, M. (2017). "Image Matching Using SIFT, SURF, BRIEF and ORB: Performance Comparison for Distorted Images", arXiv:1710.02726 [cs.CV], DOI: <https://doi.org/10.48550/arXiv.1710.02726>
- [11] Wei, X., Dong, Y., Liu, Q., Wang, L., and Lou, L. (2024). "Robust Corner Detector in Continuous Space," The Visual Computer, 1-14, DOI : <https://doi.org/10.1007/s00371-024-03362-x>
- [12] Yazdi, R. and Khotanlou, H. (2024). "Robust Corner Detector Based on Local Maximum and Minimum Differences," the 10th International Conference on Web Research, Tehran, Iran, 92-98, DOI: <https://doi.org/10.1109/ICWR61162.2024.10533379>.
- [13] Tatit, Pornrawee, Kiki Adhinugraha, and David Taniar. (2024). "Navigating the Maps: Euclidean vs. Road Network Distances in Spatial Queries" Algorithms, 17, 1-29, DOI: <https://doi.org/10.3390/a17010029>
- [14] Parkavi, A., Alex, S.A., Pushpalatha, M.N. A. Shukla, P.K., Pandey, A. and Sharma, S. (2024). "Drone-Based Intelligent System for Social Distancing Compliance Using YOLOv5 and YOLOv6 with Euclidean Distance Metric," SN Computer Science, 5, 972. <https://doi.org/10.1007/s42979-024-03304-3>
- [15] Moon, Y.R., Son, G.E., Nam, G.U., and Lee, H.J. (2024). "Research on Constructing a Sentiment Lexicon for the F&B Sector based on the N-gram Framework," Journal of The Korea Society of Computer and Information, 29(10), 11-19. DOI: <https://doi.org/10.9708/jksci.2024.29.10.011>

Authors



Seo-Hyuck Lee is currently a Senior Researcher at WEVEN Corp., after graduating from the Department of Chemistry and Medical Chemistry at Yonsei University. He have led the development of vision AI

programming and firmware for container-type smart farms, and developed VR programs specifically designed to assist patients with mild cognitive impairment. His primary research focus is in AI, with particular interests in automation and welfare technology.



Jin-san Kim is currently working as a Senior Researcher at Weven Corp. (The Global AI Start-up) after graduating from the Seoul Cyber University with a Bachelor's degree in Big Data and Information Security.

His research specialization includes AI & Computer Vision. Kim received the Gold Prize at the ICT AI Innovation competition and awarded the Best Project and Outstanding Student Award in the Multicampus AI Program. His research focuses on AI system optimization, particularly in improving computational efficiency in deep learning models.



Jung-Hwan Kim received the B.S. degree in Management from Korea University in 2009 and Ph.D. degrees in Digital Management from Korea University, Korea in 2024 respectively.

Dr. Kim opened the cutting edge business in Generative AI powered solution, as named "WEVEN(Corp.)". He is currently the founder and CEO of the Weven Corp. since 2022. He is interested in Deep Learning, Artificial Intelligence, Automatic Process Engineering Start-ups and Entrepreneurship.



Hanjin Lee received the B.S. in Sociology, M.S. degree in Communication from Yonsei University, Korea in 2006 and 2011. Furthermore He received the Ph.D. degrees in Digital Management from Korea University,

Korea, in 2021 honorably. Dr. Lee's research interests are focused on AI biz innovation, as well as the intersection between user behavior and digitalization. He has gained enormous experience in the e-Commerce industry (Naver, eBay, and Coupang Corp.)