AI 언어 모델(VLM) 기반 도시 보행 단절 구간의 개선 방안 제시

서울시 종로구 일대를 중심으로

2020860013 김찬호
2023875030 황수진

INTRO

1. 연구 배경

최근 도시 보행 환경 패러다임은 차량 통행 위주 설계에서 사람을 최우선으로 하는 보행자 중심으로 전환되고 있다. 이러한 기조에 발맞춰 보행 공간의 질적 향상을 도모하기 위한 각종 정책과 학술 연구, 도보환경 개선 사업이 다각도로 전개되는 추세다.

학위논문
보행자 중심의 보행환경개선에 관한 연구 : (완주군 삼례후정 동학지구 중심으로)
A study on improvement plans of pedestrian environment
한영천 (전북대학교 환경대학원 환경계획학(도시및지역계획) 국내석사)

'… 보행환경개선을 추진해 나갈 때 보행자들이 안심하고 보행할 수 있는 보행의 자유, 쾌적한 보행이 이루어져 보행활성화를 달성할 수 있을 것이다.'

다수의 국내외 선행 연구는 보행환경의 질적 개선이 실질적인 보행 활성화로 직결된다는 점을 실증적으로 보고하고 있다

맞춤형 공간 개선안 / **한정된 자원의 효율적 분배** / **정책 수립을 위한 법적 당위성 확보절차**

도시의 보행 특성과 문제는 지역마다 다르기에, 획일적인 정비보다 '도보환경평가'를 통한 정확한 현상 진단이 선행되어야만 맥락에 맞는 맞춤형 공간 처방을 도출할 수 있다. 또한 한정된 공공예산, 행정력을 효율적으로 운영하기 위해서는 객관적 평가를 바탕으로 시급성과 투자 효과를 고려한 과학적 우선순위 설정이 필수적이다. 마지막으로,「보행안전법」에 근거한 5년 주기의 실태조사와 평가는 성공적인 정책 수립과 법적 당위성을 확보하기 위해 반드시 거쳐야 하는 핵심 절차이다

도보환경평가방법 기반 도보환경 개선사업 진행 필요

기존 보행환경 평가 방식

한계 1 | segmentation 기반 보행환경 평가의 한계

딥러닝 기반 평가는 이미지 내 픽셀의 물리적 비율 측정에 그쳐, 실제 보행자가 체감하는 억압감, 유쾌함 등 심리적·감성적 요소를 반영하지 못하는 치명적 한계가 존재함.

인지적 요소의 한계들

한계 2 | 선행연구(설문+AI 훈련)의 시공간적 한계

인지적 요소를 반영하기 위해 설문 및 현장 관찰 데이터를 결합한 딥러닝 예측(RSS-CNN 등)이 시도되었으나, 데이터 수집에 막대한 시간과 비용이 요구되어 광범위한 도시 단위로 분석을 확장하는 데 뚜렷한 제약이 존재함.

한계 3 | 지역적 맥락(Context) 정량화의 실패

한국 특유의 골목길 정취나 보차혼용도로에서의 불안감 같은 복합적이고 장소 중심적인 느낌을 전국 단위의 일관된 데이터로 수집하고 수치화하는 데 실패함.

도보환경평가를 심층적으로 분석하여 구체적인 보행환경 개선 방안을 도출해 내는 체계적인 접근이 필요함.

2. 연구 목적

기존 segmentation + 설문조사 기법 → VLM (시각언어모델) 도입

- 보행자의 '심리적, 인지적 느낌' 등 넓은 지역에 걸쳐 정확하게 점수화하는 새로운 평가 기준
- 새로운 분석법 기반 종로구 내 보행문제 지역 발굴 및 공간 맞춤형 설계대안 제시

1. 새로운 보행환경 평가 '방법론' 제시

기존 segmentation과 설문조사 기법의 한계를 극복하기 위해, VLM(시각언어모델) 기반 방법론을 제시한다. 해당 방법은 사람의 '심리적, 인지적 느낌'까지 넓은 지역에 걸쳐 정확하게 점수화하는 새로운 평가 기준이다.

2. 도보 취약 구간 발굴 및 '맞춤형 공간 개선안' 도출

VLM 기반의 분석법을 지역적 맥락이 잘 녹여져 있는 종로구를 대상으로 삼아, 해당 지역 내 숨겨진 보행 문제 지역을 찾아낸다. 해당 보행 문제 지역의 특성 및 장소성을 분석하고, 이를 잘 반영하여 구체적인 맞춤형 공간 개선안을 제안한다.

SITE

대상지 – 서울특별시 종로구 일대

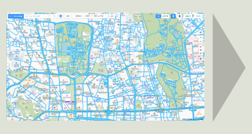
역사 문화 자원과 상업, 주거 시설이 혼재되어 있는 서울특별시 종로구 일대 (북촌, 서촌, 경복궁 등) 를 대상지로 설정하였다. 전통 한옥과 좁은 골목, 현대 대로가 공존해 일상보행과 여가보행이 다채롭게 발생한다는 특성을 가진다.

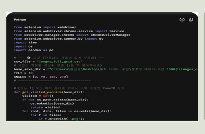
북촌 | 한옥마을 기반 여가보행 특화 구역
서촌 | 좁고 불규칙한 골목길 기반 상업 시설과 거주지가 복잡하게 얽혀 있는 구역
청계천 및 역사 인근| 현대와 과거가 혼재되어 있고 물리적 보행 폭은 넓지만 차량 통행량이 많아 심리적 보행 환경의 편차가 큰 구역

PROCESS

1. 도보환경 데이터셋 수집 환경 구축 및 진행

(1) 도보환경 이미지 데이터셋 수집


카카오맵 기반 API


파이썬 기반 자동화 코드 실행

노드당 0, 90, 180, 270도 전방향 이미지 추출

가장 많은 자료를 제공하는 카카오맵 API를 본 연구의 도보환경 이미지 데이터 수집 맵으로 선정했다. 이후 파이썬 기반 카카오맵 API 내 이미지 추출 자동화 코드를 구축해 30m 단위마다 노드를 추출하고, 한 노드에서 360 º 전방향 이미지를 수집해 데이터셋을 완성했다.

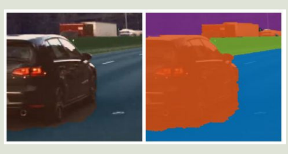
(2) 데이터 대상 및 수집량

종로구 동 87개 中 80개 대상 데이터 수집 → 총 추출 노드 9,200개 추출 전방향 이미지 약 36,700장

* 제외 동 : 평창동, 구기동, 부암동, 홍지동, 신영동, 창신동, 숭인동

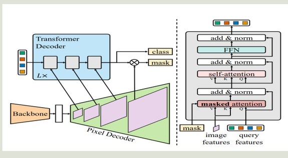
2. Mask2Former 기반 image segmentation 도보환경 평가 진행

(1) image segmentation 소개 및 선정 모델



이미지 세그멘테이션(Image Segmentation)

픽셀 단위 분류를 통해 이미지를 분석하는 기법이다. 보행로나 장애물 존재 인식을 넘어, 실제 보행 가능한 유효 폭, 단차나 적치물의 정확한 경계, 녹지 등이 시야에서 차지하는 비율 등을 픽셀 수준에서 계산한다. 즉 물리적 보행환경을 수치 요소로 변환하여 보행의 안전성과 쾌적성을 공간적으로 평가할 수 있는 핵심적인 근거를 제공한다.



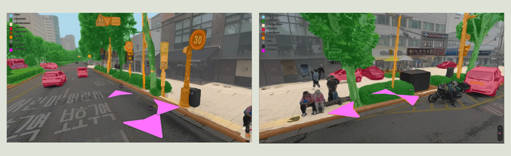
Mask2Former 모델

대다수의 모델은 주로 서구권의 정형화된 보행환경을 학습하여, 보차혼용도로나 좁은 이면도로, 특유 적치물 등이 혼재되어 있는 아시아의 복잡한 도심 환경을 인식하는 데 한계를 지닌다. 반면, 아시아 도보환경에 최적화된 본 모델은 대상지인 종로구 현장 상황을 가장 객관적으로 지표를 산출하게 하여 선정했다. 또한 모델의 validation mIoU는 약 83%로, 신뢰 가능함을 확인했다.

(2) image segmentation 분류 클래스 및 진행과정

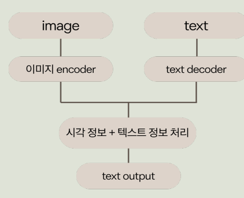
선행연구를 기반으로 아래와 같은 클래스를 선정하였으며 클래스별 픽셀수를 기반으로 수치화, (positive class)-(negatvie class)로 점수화를 진행했다.

positive class	negative class	etc
vegetation, sky, sidewalk	vehicle, wall, fence, obstacle	building, arrow, road



3. Google Gemini 기반 VLM 도보환경 평가 진행

(1) VLM 소개 및 선정 모델



Vision Language Model (시각언어모델)

시각 정보(이미지, 비디오)와 언어 정보(텍스트)를 동시에 이해하고 처리할 수 있는 멀티모달(Multimodal) 인공지능 모델로, 흔히 chat-gpt, claude ai, gemini 등이 있다.

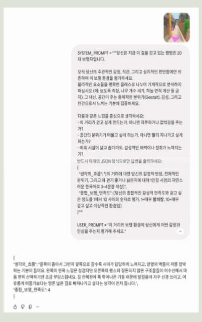
본 연구에서는 google Gemini-Flash 3.0 모델을 사용했다.

(2) VLM 모델 신뢰도 검증 과정

본 연구에서 제안하는 VLM 기반 정성 평가 방법론의 타당성과 객관성을 확보하기 위해, 랜덤 추출된 VLM output 이미지 100장을 대상으로 인간 평가자와 VLM 간의 교차 검증을 수행하였다. 그 결과, 82%의 높은 상호 일치도를 보임을 확인했고, 본 평가 모델이 신뢰, 일관성 있게 평가할 수 있는 유효한 방법론임을 증명했다.

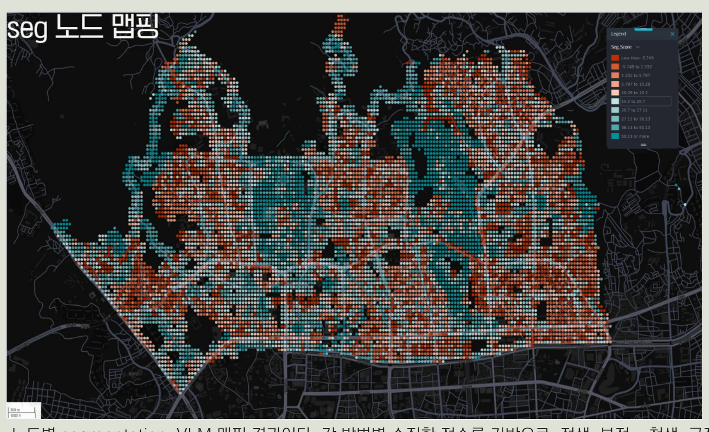
(3) VLM 기반 보행환경 평가 진행과정

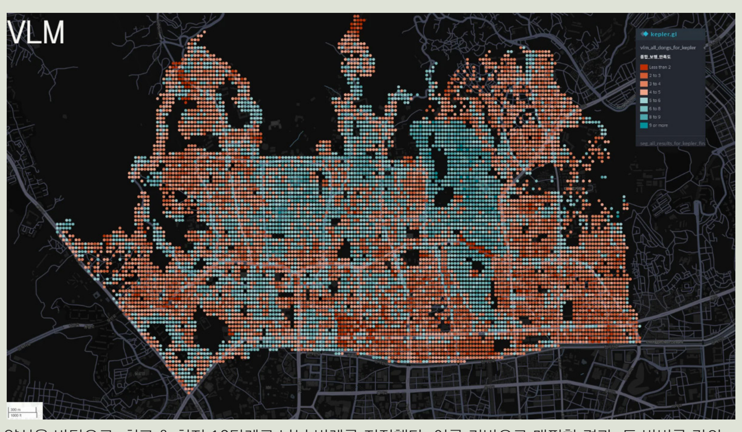
앞서 각 노드에서 추출한 전방향 이미지, 선행연구 기반 텍스트 프롬프트를 VLM에 제공하여 VLM 기반 보행환경 평가를 진행했다. 모델을 20대 일반 보행자로 설정해 사진 속 보행환경을 주관적 감정, 직관, 보행쾌적성에 의존해 판단하도록 했다. 또한, 생각의 흐름과 종합보행점수(1~10점 범위)를 제시하도록 하여 이를 바탕으로 VLM 기반 보행환경 평가 결과를 얻었다.



4. 방법간 교차 분석

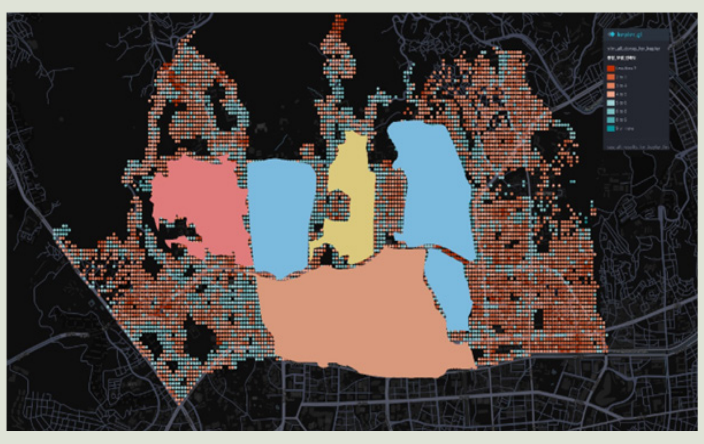
(1) 방법별 매핑 결과

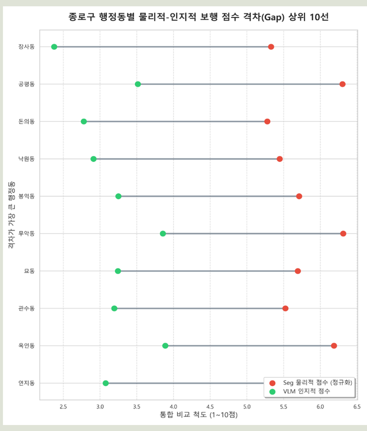


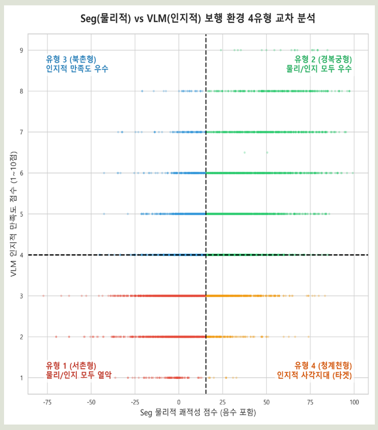


노드별 segmentation, VLM 매핑 결과이다. 각 방법별 수집한 점수를 기반으로, 적색-부정 ~ 청색-긍정 양상을 바탕으로, 최고 & 최저 10단계로 나눠 범례를 지정했다. 이를 기반으로 매핑한 결과, 두 방법론 간의 비슷한 결과가 보이는 구간이 있었지만 동시에 차이(갭)가 발생하는 구간이 있음을 확인할 수 있었다.

(2) 노드 매핑 기반 구역화 및 시각화





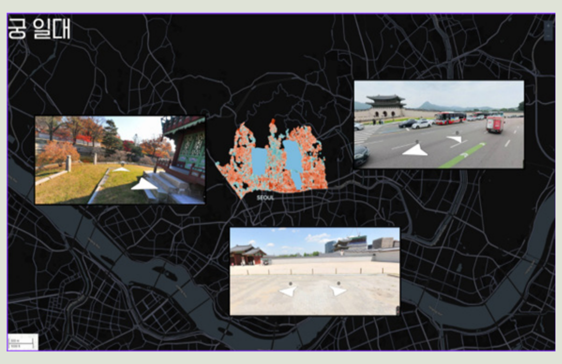


앞서 언급했던 데이터 기반으로 통계를 내었을 때 발생하는 갭 차이 구간을 바탕으로, 위 사진과 같이 도식화하여 다음과 같이 적색, 청색, 황색, 주황색 구간으로 나눌 수 있었다. 나아가, 해당 갭 차이 발생구간에 있어 노드 결과 값을 바탕으로 위와 같은 그래프 또한 얻을 수 있었다. 해당 사진들은 두 방법 간의 분석에 따라 엄연한 차이가 발생함을 보여주고 있다. 또한, 이 결과 차이를 바탕으로 그래프를 그렸을 때, 격차가 발생함을 시각화할 수 있었고, 이를 바탕으로 유형을 총 4가지로 나눌 수 있었다.

ANAYLSIS

1. 교차 분석 기반 공간별 분석 결과

서촌(유형 1) : 좁은 골목길에 차량 통행이 혼재되어 있으며, 건물 외벽에 설치된 에어컨 실외기, 입간판 등 무분별한 생활 적치물과 낡은 외벽이 보행을 방해한다. 이러한 환경은 시각적·물리적 피로를 극도로 유발하여 기계(Seg)와 인간(VLM) 모두 걷기 나쁜 부정적 환경으로 평가하였다.



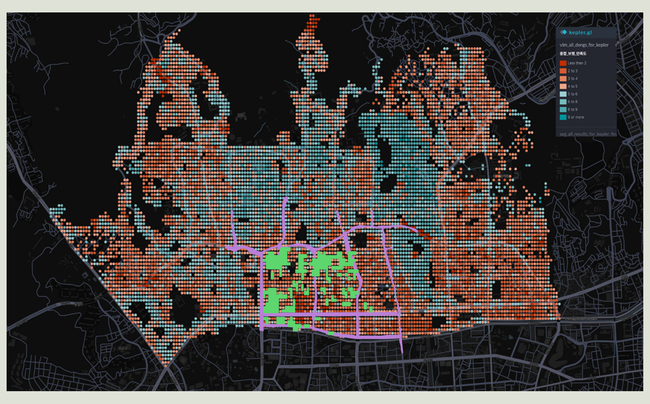
궁 일대(유형 2) : 넓은 보행 폭이 안정적으로 확보되어 있고 차도와 보도가 물리적으로 명확히 분리되어 있다. 특히 전통을 나타내는 정갈한 식재와 깔끔한 마감재가 사용되어, 기계적 수치뿐만 아니라 VLM(가상 보행자)에게도 절대적인 쾌적함과 뚜렷한 장소성을 제공하는 가장 이상적인 공간이다.

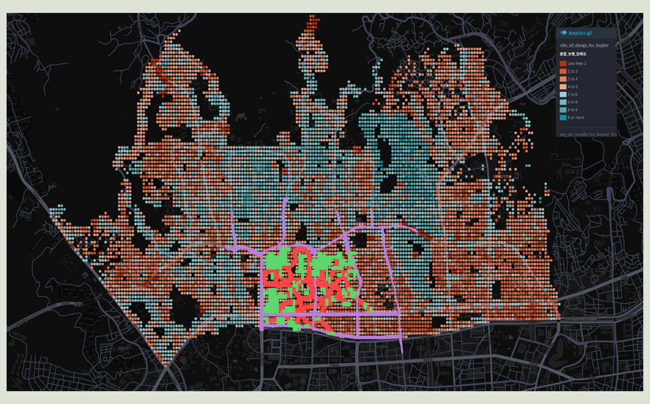
북촌(유형 3) : 길 너비 등 물리적 수치는 서촌(유형 1)과 비슷하게 좁아 Seg 점수는 낮게 산출된다. 하지만 차량 진입이 통제되고 바닥 포장이 잘 정비되어 있으며, 한옥과 돌담이 주는 고즈넉한 전통적 장소성 덕분에 보행자(VLM)는 심리적인 아늑함과 높은 만족도를 느끼는 플러스 갭 구간이다.

역 주변 일대(유형 4) : 하늘 픽셀이 넓게 열려 있고 도보 부분의 픽셀이 넓어 기계적인 Seg 점수는 높게 산출된다. 그러나 실제로는 삭막한 외벽, 장소성의 결여, 보차 혼용으로 인한 어수선함에 인간은 심리적 불안감을 느끼며 걷고 싶지 않은 길로 평가하는 마이너스 갭구간이다.

2. 교차 분석 기반 분석 결과

(1) 타겟 구역 선정 및 공간 개선 전략



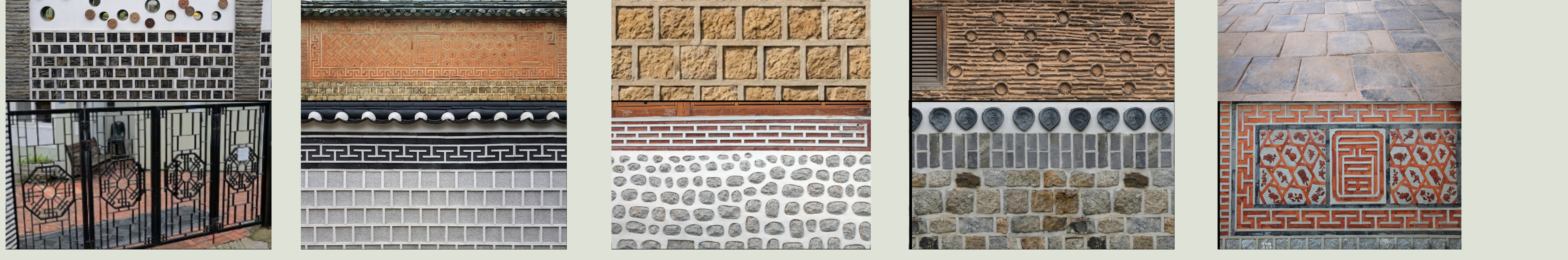


1번 유형(서촌 등)은 물리적 인프라 자체가 열악하여 기존의 물리적 분석으로도 쉽게 문제점을 발견할 수 있는 1차원적 취약지구다. 반면, 4번 유형은 도로는 넓고 시야는 트여 있어 기존의 물리적 평가 기준으로는 우수하다고 판별되어 행정적 방치 상태에 놓이기 쉬운 공간이다. 본 연구는 VLM을 통해서만 발견할 수 있는 이러한 인지적 사각지대를 타겟으로 삼아 설계적 대안을 제시함으로써, 새로운 도보환경평가 방법론의 실효성을 극대화하고자 한다.

(2) 공간별 특징 분석 및 추출

서촌(유형 1)의 부정적 요소 제거: 시각적 혼란을 야기하는 무질서한 시설물과 보차 혼용의 위협 등 보행을 꺼리게 만드는 불안 요소를 물리적으로 제거한다.
북촌·경복궁(유형 2, 3)의 긍정적 인지 요소 이식: 인위적이고 이질적인 대형 조형물 설치를 지양하고, 해당 구역이 가진 전통적 포장 질감(사고석 등)과 고유의 패턴을 유형 4 공간의 바닥재 및 마감재에 적용하여 상실된 장소성을 부여한다.

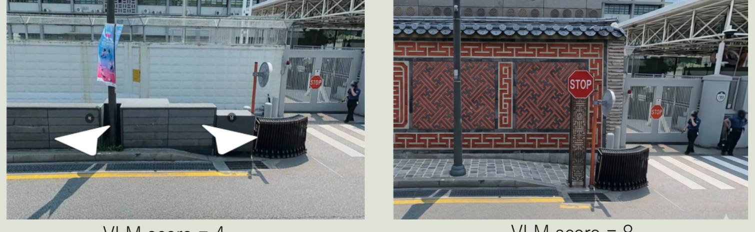
아래 사진은 북촌과 경복궁 로드뷰에서 추출한 패턴 디자인 샘플이다.



RESULT

1. 공간별 대안 제시

(1) 가로변 공간 (대로 연접 차량 혼재 구간) : 서촌 바탕 부정적 요소 제거



VLM score = 4 / VLM score = 8

차로 폭을 축소하고 잉여 공간을 보도화하는 도로 정리를 시행했다. 보차 혼용으로 인한 심리적 압박감과 무질서한 노상 적재물을 전면 정리하고, 명확한 단차 조정 및 펜스를 통해 시각적 어수선함을 정돈하였다.

(2) 이면도로 및 골목길 공간 : 북촌 내 인지적 점수 향상 요소 이식 (장소성 보완)

VLM score = 3 / VLM score = 7

기존의 평범하고 어수선한 보도블록을 걷어내고 북촌에서 차용한 전통적 질감의 블록 포장을 전면 적용하여 공간의 장소성을 개선하였고, 이를 통해 정취 있는 골목길 공간으로 공간을 개선하였다.

(3) 우수 거점 공간: 단절 구간 연결을 위한 거점 구간

VLM score = 7 / VLM score = 8

경복궁 일대에서 볼 수 있는 정갈한 한국적 패턴과 식재 방식을 보도포장과 조경에 확장 적용하였다. 흩어져 있던 양호한 보행 노드들을 하나의 시각적 흐름으로 묶어주는 거점 역할을 부여하여 단절된 구간을 연결하는 구간을 수행한다.

2. 개선된 노드 기반 가상 동선 제시

3. 기대효과 및 한계

(1) 결론 및 기대효과

본 연구는 도보환경평가에 있어 객체 인식에 머물러 있던 기존 시맨틱 세그멘테이션(물리적 수치)의 한계를 극복하기 위해 시각언어모델(VLM)을 '가상 보행자'로 도입한 융합 평가 방법론을 제시하였다. 종로구 일대 3.6만 장의 스트리트 뷰 이미지를 분석한 결과, 보도폭이나 개방감이 수치상 우수하더라도 보차 미분리나 삭막한 미관으로 인해 인지적 만족도가 낮게 나타나는 '유형 4 (인지적 사각지대)'를 발견할 수 있었다. 본 연구는 이러한 교차 분석 결과를 바탕으로, 맹목적인 물리적 인프라 확충에서 벗어나 각 가로의 맥락적 결핍을 채워주는 타겟팅된 설계 대안을 제시함으로써 진정한 보행자 중심의 공간 디자인이 나아가야 할 새로운 방향성을 입증하였다.

(2) 한계점 및 제언

본 연구가 지니는 한계점은 다음과 같으며, 향후 이를 보완하는 후속 연구가 요구된다.
시각 데이터 의존의 한계: 스트리트 뷰 이미지(GSV) 기반의 평가는 시각적 환경을 매우 훌륭하게 모사하지만, 실제 보행 환경에서 지대한 영향을 미치는 소음(차량 소음), 후각(매연, 악취), 그리고 미기후(여름철 폭염, 빌딩풍 등)와 같은 비시각적 감각 데이터를 온전히 반영하지 못하는 한계가 있다.
단일 시점 스냅샷의 한계: 수집된 이미지는 특정 계절과 시간대에 촬영된 정적인 화면이므로, 출퇴근 시간대의 극심한 혼잡도 변화나 계절의 동적인 흐름을 유연하게 반영하기 어렵다.

AI 프록시의 제약: VLM이 고도화된 형태 심리를 모사하여 인간과 90% 이상 일치하는 신뢰도를 보여주었으나, 휠체어 이용자, 시각 장애인, 영유아 동반자 등 보다 세분화되고 특수한 신체적 조건을 가진 보행자의 실제 감각을 100% 완벽하게 대체하기에는 무리가 있다.

또한 본 연구의 방법론과 도출된 결과는 향후 서울시가 추진하는「걷는 도시, 서울」등 보행 정책 수립 과정에 다음과 같이 기여할 수 있다.

보행환경 개선 예산의 효율적 배분 기준 제공: 공공 예산은 한정되어 있으므로, 단순히 물리적 수치(보도폭 등)가 좁은 곳을 일괄 정비하는 방식을 지양해야 한다. 본 연구에서 제시한 VLM-Seg 교차 분석 맵을 활용하면, 도로는 넓지만 인지적 만족도가 최악인 '방치된 취약 지구(유형 4)'를 우선적으로 솎아내어 적은 예산(포장재 변경, 펜스 설치 등)으로 가장 극적인 체감 효과를 끌어낼 수 있다.

여가보행 특화 가로 가이드라인 수립: 효율적 이동을 위한 일상보행로와 달리, 여가보행로는 '보행의 정취와 심리적 안정감'이 필수적이다. 향후 서울시의 보행특구 및 특화거리 조성 사업 시, 본 연구의 VLM 인지 지표를 서울시 가로 설계 매뉴얼의 보조 지표로 편입시킨다면 기계적인 수치를 넘어 시민의 감성과 장소성을 채워주는 진정한 의미의 '보행특별시'를 완성하는 강력한 정책적 근거가 될 것이다.