

Try-Angle AI: AI 기반 실시간 사진 구도 가이드 및 미적 평가 시스템

중앙대학교 예술공학부 캡스톤 디자인 프로젝트 1

프로젝트 기획서

팀명 : Try-Angle

팀원 : 김현수, 김세영, 이윤균, 전은서, 최승혜

1. 프로젝트 개요

1.1 배경 및 필요성

현대 사회에서 사진은 단순한 기록을 넘어 자기표현과 소통의 핵심 수단이 되었습니다. 특히 SNS의 확산으로 일상적인 사진 촬영이 보편화되었지만, 대부분의 사용자는 전문적인 사진 교육을 받지 못해 좋은 구도를 찾는 데 어려움을 겪고 있습니다. 기존 카메라 앱들은 단순한 격자선이나 수평 가이드만을 제공할 뿐, 실제 구도 개선을 위한 능동적인 피드백은 제공하지 못하고 있는 상황입니다.

저희 프로젝트는 딥러닝 기반의 객체 검출 및 포즈 분석 기술을 활용하여, 촬영 중 실시간으로 구도를 분석하고 즉각적인 개선 방안을 제시하는 혁신적인 촬영 보조 시스템을 개발하려고 합니다. 이를 통해 사진 초보자도 전문가 수준의 구도를 쉽게 구현할 수 있도록 지원하며, 촬영 후에는 미학적 평가를 통해 지속적인 학습과 개선이 가능하도록 합니다.

저희 프로젝트는 일상생활에서 사진을 잘 찍지 못하는 사람들을 위한 카메라 앱 개발이며, 사진학적 구도를 학습하고 전문가 수준의 퀄리티를 원하는 전문가의 관점은 아직 다루지 않을 예정입니다.(추후 개발 예정입니다)

1.2 프로젝트 목표

Try-Angle AI는 다음과 같은 핵심 목표를 달성할 계획입니다.

실시간 구도 가이드 제공

촬영 중 카메라 뷰파인더에서 객체 검출 및 구도 분석을 수행하여 즉각적인 피드백을 제공함으로써, 사용자가 촬영 전에 최적의 구도를 찾을 수 있도록 지원합니다.

전문가 수준의 사진학 이론 적용

삼분할 법칙, 황금비율, 리딩 라인 등 전통적인 사진학 이론을 알고리즘화하여, 이론적 근거를 바탕으로 한 정확한 구도 평가를 제공합니다.

SNS 트렌드 기반 맞춤형 제안

Instagram, Pinterest 등 SNS 플랫폼의 인기 사진 패턴을 분석하여, 현대적 감각과 트렌드를 반영한 구도 추천을 제공합니다.

촬영 후 미학적 평가 시스템

NIMA(Neural Image Assessment) 모델을 활용하여 촬영된 사진의 미학적 품질을 정량적으로 평가하고, 구체적인 개선 방안을 제시합니다.

1.3 프로젝트의 특징

실시간성(Real-time Feedback)

기존의 사진 앱들은 사진 촬영 전/후의 보정만을 진행해줄 뿐 사진 구도에 대한 피드백이나 포즈에 대한 피드백은 제공해주지 않습니다. 저희 프로젝트는 촬영 후 편집이 아닌 촬영 중 즉각적인 구도 피드백을 제공함으로써, 사용자가 최적의 순간을 놓치지 않고 포착할 수 있습니다. 평균 60~80fps(아이폰 15pro 기준)을 목표로 하고 있으며 자연스러운 촬영 경험을 보장합니다.

정확성(Accuracy)

최신 포즈 및 객체 검출 모델인 YOLOv11의 pose 및 segmentation을 활용하여 사람을 90%이상의 객체 검출 정확도를 달성하며, 사진학 이론과 통계적 분석을 결합한 과학적 구도 평가를 제공합니다.

실용성(Practicality)

단순히 전통적 사진 구도 및 이론에 기반한 가이드가 아닌, SNS에서 인기있는 트렌드 구도를 학습하여 예술적 가치와 대중적 선호도를 모두 고려한 실용적인 가이드를 제공합니다.

접근성(Accessibility)

전문 용어를 배제하고 일상적인 언어로 피드백을 제공하며, 사진을 찍는 동안 직관적인 가이드를 시각적 및 언어적으로 동시에 제공해줍니다.

2. 기존 솔루션 분석 및 차별점

2.1 기존 솔루션의 한계

현재 시장에는 다양한 카메라 앱과 사진 편집 도구가 존재하지만, 실시간 구도 가이드 측면에서는 뚜렷한 한계를 보이고 있습니다.

기본 카메라 앱은 대부분 격자선과 수평 가이드 정도의 단순한 보조선만을 제공합니다. 이는 사용자가 이미 구도 이론을 알고 있다는 전제 하에 수동적인 참고자료를 제공하는 것에 불과하며, 현재 구도의 문제점이나 개선 방법을 능동적으로 알려주지 않습니다.

따라서 초보자는 격자선이 무엇을 의미하는지 이해하지 못한 채 그대로 지나치는 경우가 많습니다.

전문 카메라 앱은 수동 설정 기능과 다양한 필터를 제공하지만, 오히려 이러한 복잡한 기능들이 초보자에게는 진입장벽으로 작용합니다. 또한 이러한 앱들 역시 실시간 구도 안내 기능은 제공하지 않으며, 촬영 기술보다는 후보정에 중점을 둡니다.

사진 편집 앱은 강력한 후보정 기능을 제공하지만, 촬영 시점의 도움은 전혀 제공하지 않습니다. 이미 촬영된 사진의 구도 문제는 편집으로 해결하기 어려운 경우가 많으며, 재촬영이 불가능한 순간의 기회를 놓치게 됩니다.

AI 미학 평가 서비스는 촬영된 사진에 대해 점수를 제공하지만, 실시간 피드백이 없고 구체적인 개선 방법 제시가 미흡합니다. 단순히 "구도가 좋지 않습니다"라는 평가만으로는 사용자가 다음 촬영에서 무엇을 개선해야 하는지 알기 어렵습니다.

2.2 본 프로젝트의 차별점

Try-Angle AI 는 다음과 같은 핵심 차별점을 통해 기존 프로젝트와의 차별성을 제공할 예정입니다.

실시간 구도 분석 및 즉각적 행동 유도

촬영 중 매 프레임을 분석하여 "오른쪽으로 2 발자국 이동하세요", "피사체가 배경 기둥과 겹칩니다" 등 구체적이고 즉각 실행 가능한 피드백을 제공합니다. 이는 사용자가 스스로 문제를 파악하고 해결책을 찾아야 하는 기존 방식과 달리, AI 가 능동적으로 가이드하는 새로운 패러다임입니다.

사진학 이론과 SNS 트렌드의 융합

삼분할 법칙, 황금비율 등 검증된 사진학 이론을 기반으로 하되, Instagram 과 Pinterest 의 인기 사진 패턴 분석을 통해 현대적 감각을 더했습니다. 사용자는 클래식 모드와 트렌드 모드를 선택하여 자신의 촬영 목적에 맞는 가이드를 받을 수 있습니다.

촬영 전 가이드와 촬영 후 평가의 통합

촬영 중에는 yolo 기반의 실시간 구도 가이드를, 촬영 후에는 NIMA 기반 미학 평가를 제공하는 완전한 순환 학습 시스템입니다. 사용자는 즉각적인 피드백으로 좋은 사진을 얻고, 평가를 통해 자신의 촬영 실력을 지속적으로 향상시킬 수 있습니다.

모바일 온디바이스 AI 구현

클라우드 연결 없이 스마트폰에서 직접 AI 추론을 수행하여, 네트워크 지연 없는 빠른 반응속도(yolov11 모델 사용 시 예상 FPS: 70~100fps)와 프라이버시 보호를 동시에 달성합니다. 사용자의 사진이 외부 서버로 전송되지 않아 개인정보 보호 측면에서도 안전합니다.

3. 기술 구현 방안

3.1 설계 구조

본 시스템은 크게 실시간 프레임 처리, AI 기반 분석, 피드백 생성, 후처리 평가의 4 단계 파이프라인으로 구성됩니다.

사용자가 카메라를 실행하면, 시스템은 초당 40 ~ 60 프레임의 실시간 영상 스트림을 캡처합니다. 각 프레임은 모바일 디바이스에 탑재된 경량화 AI 모델을 거쳐 분석되며, 이 과정에서 인물 및 배경 객체의 위치, 신체 자세, 구도 규칙 준수 여부 등이 종합적으로 평가됩니다. 분석 결과는 즉시 피드백 생성 모델(자체 알고리즘 개발)로 전달되어, 사용자 화면에 시각적 가이드라인과 텍스트 안내로 표시됩니다. 촬영이 완료되면, 후처리 평가 시스템이 작동하여 NIMA 모델 기반의 미학적 점수를 산출하고 개선 제안을 제공합니다.

3.2 핵심 AI 모델

객체 검출 : YOLOv11-small

YOLOv11-small 모델은 YOLOv11의 경량 버전으로, 실시간 객체 검출과 인체 포즈, 객체 segmentation 분석에 최적화되어 있습니다. 이 모델을 이용해 프레임 내 인물과 주요 배경 객체(나무, 건물, 벤치 등)의 위치를 정확히 파악하며, 인물 검출과 비교해도 90% 이상의 정확도 및 다양한 객체의 위치와 영역 정보를 동시에 산출할 수 있습니다. 바운딩 박스와 segmentation 마스크, 신체 keypoint 를 모두 활용해 삼분할 법칙 준수 여부, 객체 간 겹침, 구도 평가 등을 자동으로 판단합니다.

YOLOv11은 객체 검출뿐만 아니라 내장된 pose 및 segmentation 기능을 통해 인물의 자세와 구성, 프레임 내 모든 객체의 영역을 실시간으로 분석합니다. 별도의 모델(SAM2, RTM-POSE, MediaPipe 등)을 쓰지 않고, YOLOv11 단일 모델 아키텍처만으로 실시간 분석이 가능합니다.

모델은 TensorFlow Lite(Android) 및 Core ML(iOS) 포맷으로 변환되어 온디바이스에서 직접 추론하며, 양자화(quantization)를 적용해 모델 크기와 연산량을 줄이면서도 정확도 손실을 최소화합니다.

포즈 분석 YOLOv11 – small – pose/seg

YOLOv11의 pose 기능을 통해 33 개 신체 keypoint 와 신체 영역(segmentation)을 동시에 추출하여 인물 자세와 균형, 팔·다리 위치 등 세부 포즈를 정밀 분석합니다. 어깨 각도나 신체 배치 등이 부자연스러운 경우 "자세를 조금 더 펴보세요" 등의 실시간 피드백을 제공하며, 신체 일부가 프레임 밖으로 벗어나 있을 경우 "카메라를 조금 뒤로 이동하세요"와 같은 안내를 자동으로 띄울 예정입니다.

미학 평가: NIMA(Neural Image Assessment)

저희는 Google Research에서 개발한 신경망 기반 이미지 미학 평가 모델인 NIMA를 통해 구도, 색감, 밝기, 대비 등 다양한 미학적 요소를 평가 후 피드백을 진행할 예정입니다.

촬영 후 이 모델을 통해 사진의 전체적인 품질을 객관적으로 평가하고, 항목별 세부 점수를 제시하여 사용자가 어느 부분을 개선해야 하는지 명확히 파악할 수 있도록 합니다.

3.3 구도 평가 알고리즘

구도 평가 알고리즘은 전통적인 사진학 이론을 알고리즘을 만든 규칙 기반(rule-based) 시스템입니다.

삼분할 법칙

프레임을 가로세로 3등분하여 생성된 그리드의 교차점 근처에 주 피사체가 위치하는지 평가합니다. 피사체의 중심점과 가장 가까운 교차점 간의 거리를 계산하여, 임계값 이내일 경우 높은 점수를 부여합니다. 이 항목은 전체 평가에서 30%의 가중치를 차지합니다.

겹침 방지

인물 바운딩 박스와 배경 객체 바운딩 박스의 IoU(Intersection over Union)를 계산하여, 5% 미만일 경우 이상적으로 판단합니다. 겹침이 발생할 경우 "피사체를 왼쪽으로 이동하세요" 등의 구체적인 안내를 제공합니다.

프레이밍

주 피사체가 프레임 중심에서 적절한 거리를 유지하는지 평가합니다. 너무 가까우면 답답하고, 너무 멀면 피사체가 작아 보이는 문제를 방지하기 위해, 피사체가 프레임 면적을 차지하는 비율을 고려해 개발할 예정입니다.

포즈의 자연스러움

YOLO로 추출한 keypoint를 기반으로 어깨 각도, 팔다리 배치의 대칭성, 신체 비율 등을 분석합니다. 전문 사진작가가 선호하는 포즈 패턴을 학습하여, 이와 유사할수록 높은 점수를 부여합니다.

공간 활용

피사체 주변의 빈 공간이 적절히 배치되었는지 평가합니다. 과도하게 비어있거나 혼잡하지 않고 시각적 균형을 이루는지 판단합니다.

각 항목의 점수를 가중 평균하여 0~100 점의 최종 구도 점수를 산출하며, 70 점 이상을 권장 구도로 판단합니다.

3.4 SNS 트렌드 기반 구도 추천

전통적인 사진학 이론만으로는 현대 SNS 사용자들의 선호도를 완전히 반영하기 어렵습니다. 이에 본 프로젝트는 Instagram, Pinterest 등 주요 SNS 플랫폼의 인기 사진 데이터를 수집하고 분석하여, 통계적으로 유의미한 구도 패턴을 추출합니다.

약 1,000~10,000 장의 높은 참여도를 받은 사진(좋아요, 댓글 수 상위)을 수집하여, YOLO로 구도 feature를 추출한 뒤, 일반 사진과의 통계적 차이를 분석합니다. 예를 들어, 여행 사진에서는 피사체를 프레임 하단에 배치하고 배경을 넓게 보여주는 구도가 인기 있으며, 인물 셀카에서는 약간 위에서 아래로 촬영하는 각도가 선호되는 등의 패턴을 발견할 수 있습니다.

이러한 트렌드 규칙은 "트렌드 모드" 옵션으로 제공되며, 사용자는 클래식한 사진학 이론과 현대적 SNS 트렌드 중 원하는 스타일을 선택할 수 있습니다.

3.5 개발 환경 및 도구

개발환경은 docker를 이용하여 모델 별 환경을 다르게 구성하여 관리하려고 했지만, 여러 문제로 인해 가상환경을 별도로 만들어 관리할 예정입니다.

버전 관리는 Git과 GitHub를 통해 이루어지며, 협업 도구로는 Notion(문서화), Figma(UI/UX 디자인), discord(소통)를 사용합니다.
모바일 앱 테스트는 아이폰 개발환경인 xcode로 예상하고 있습니다.

4. 기대효과 및 활용방안

4.1 사용자 측면의 기대효과

사진 촬영 실력의 빠른 향상

기존에는 좋은 사진을 찍기 위해 사진학 서적을 읽거나 강의를 들어야 했지만, 본 시스템은 실전에서 즉각적인 피드백을 통해 자연스럽게 학습할 수 있게 합니다. 사용자는 여러 번의 촬영을 통해 자신도 모르게 좋은 구도를 학습하게 되며, 결과적으로 전문 교육 없이도 높은 수준의 사진을 촬영할 수 있습니다.

재촬영 기회 상실 방지

여행이나 특별한 순간처럼 재촬영이 불가능한 상황에서, 촬영 중 실시간 피드백은 첫 시도에서 좋은 사진을 얻을 확률을 크게 높입니다. 기존에는 집에 돌아와서야 사진이 잘 나오지 않았음을 깨닫고 후회하는 경우가 많았지만, 본 시스템은 그 자리에서 문제를 수정할 수 있게 합니다.

SNS 참여도 향상

통계적으로 검증된 인기 구도를 추천함으로써, 사용자의 SNS 게시물이 더 많은 좋아요와 댓글을 받을 가능성이 높아집니다. 이는 특히 인플루언서나 소상공인처럼 SNS가 중요한 사용자에게 큰 가치를 제공합니다.

자신감 향상

객관적인 평가 점수와 긍정적 피드백은 사용자에게 자신감을 줍니다. "나도 좋은 사진을 찍을 수 있다"는 자기 효능감은 더 적극적인 사진 촬영으로 이어지며, 이는 다시 실력 향상으로 연결되는 선순환을 만듭니다.

4.2 산업적 측면의 기대효과

사진 교육 시장의 혁신

전통적인 오프라인 사진 강의나 온라인 강좌를 보완하는 실전 학습 도구로 자리잡을 수 있습니다. 교육 기관이나 사진 커뮤니티와 제휴하여 학습 프로그램의 일부로 통합될 가능성이 있습니다.

카메라 및 스마트폰 제조사와의 협업

본 기술을 스마트폰 제조사에 라이선스화하여 기본 카메라 앱에 통합하거나, 카메라 제조사의 제품에 탑재할 수 있습니다. 이는 제품의 차별화 요소가 되어 판매 증대에 기여할 수 있습니다.

관광 및 여행 산업 활용

관광지에서 본 앱을 활용한 "포토 스팟" 가이드를 제공하거나, 여행사가 고객에게 추천하는 앱으로 포지셔닝할 수 있습니다. 사용자가 여행지에서 더 좋은 사진을 찍게 되면 SNS를 통한 자연스러운 홍보 효과도 기대할 수 있습니다.

광고 및 마케팅 분야 적용

전문 사진작가나 광고 대행사에서 클라이언트와의 커뮤니케이션 도구로 활용할 수 있습니다. "이 구도는 85 점입니다"와 같이 정량화된 지표는 주관적 평가를 객관화하는데 도움이 됩니다.

5. 결론

5.1 프로젝트 의의

Try-Angle AI는 전문가의 영역으로 여겨지던 좋은 구도와 미학적 감각을 누구나 접근할 수 있게 만들어, 일상의 순간들을 더욱 아름답게 기록할 수 있도록 돕습니다.

기술적으로는 딥러닝 기반 객체 검출, 포즈 분석, 미학 평가라는 최신 AI 기술을 실시간 모바일 환경에서 구현하는 도전적인 프로젝트라고 생각합니다. 이는 온디바이스 AI의 가능성을 보여주는 실증 사례가 되며, 향후 다양한 예술적, 창의적 영역에서 AI가 인간을 보조하는 모델의 선례가 될 것입니다.

사회적으로는 시각적 리터러시와 미적 감수성을 향상시켜, 더 풍부한 사진 문화를 만드는 데 기여합니다. 특히 디지털 격차를 줄이고, 나이나 배경에 관계없이 누구나 자신을 표현할 수 있는 도구를 제공한다는 점에서 포용적 가치를 지닙니다.

5.2 차별화된 강점

기존 솔루션들이 촬영 후 편집이나 단순 보조선 제공에 그친 반면, 저희의 프로젝트는 촬영 중 실시간 능동적 가이드를 제공하여 근본적으로 다른 접근을 취합니다. 이는 재촬영이 불가능한 순간을 포착하는 데 결정적 차이를 만듭니다.

과학적 근거와 현대적 감각의 결합도 핵심 강점입니다. 검증된 사진학 이론을 기반으로 하되, SNS 빅데이터 분석을 통해 현대인의 실제 선호도를 반영함으로써, 예술성과 대중성을 모두 잡았습니다.

온디바이스 AI 구현으로 빠른 반응속도와 프라이버시 보호를 동시에 달성한 점도 차별화 요소입니다. 사용자의 사진이 외부로 전송되지 않아 안전하며, 네트워크 없이도 모든 기능을 사용할 수 있습니다.

5.3 실현 가능성

시장 측면에서 SNS 사용이 보편화되고 사진이 소통의 핵심 수단이 된 현재, 좋은 사진을 찍고자 하는 수요는 명확히 존재합니다. 특히 인스타그램 월 활성 사용자 20억 명, 핀터레스트 4.5 억 명 등 방대한 잠재 시장이 있으며, 이들 중 상당수는 현재 만족스러운 촬영 도구를 찾지 못하고 있습니다.

기술적으로 YOLO, NIMA 등 필요한 핵심 기술은 이미 검증되었으며, 모바일 하드웨어도 온디바이스 AI를 구동하기에 충분히 발전했습니다. 남은 것은 이들을 효과적으로 통합하고 사용자 경험으로 완성하는 것으로, 이는 충분히 달성 가능한 목표라고 생각합니다.