# **과일 품질 등급 분류를 통한 소비자 의사결정 방향 제시 연구**

## **1. 서론**

### **1.1. 연구 배경**

현대 소비자들은 농산물 구매 시 품질을 중요한 의사결정 요인으로 고려한다. 특히 동네 마트와 같은 소규모 유통 채널에서는 과일의 품질 등급이 명확하게 표시되지 않는 경우가 많아, 소비자들이 육안으로만 품질을 판단해야 하는 어려움이 존재한다. 이러한 정보 비대칭은 소비자의 합리적 구매 의사결정을 저해하며, 품질에 대한 불확실성은 구매 만족도 저하로 이어질 수 있다.

최근 컴퓨터 비전과 딥러닝 기술의 발전은 농산물 품질 분류 자동화의 가능성을 열었다. 특히 객체 탐지(Object Detection) 기술은 이미지에서 과일의 종류를 식별하는 동시에 품질 등급을 분류할 수 있는 잠재력을 지니고 있다. 이는 소비자에게 객관적인 품질 정보를 제공함으로써 보다 합리적인 구매 결정을 지원할 수 있다.

### **1.2. 연구 목적**

본 연구는 딥러닝 기반 객체 탐지 모델을 활용하여 과일의 품질 등급을 자동으로 분류하고, 이를 통해 동네 마트를 이용하는 소비자의 의사결정을 지원하는 시스템을 구축하는 것을 목적으로 한다. 구체적으로 YOLOv5와 EfficientDet 모델의 성능을 비교 분석하여 과일 품질 분류에 가장 적합한 모델을 도출하고자 한다.

### **1.3. 연구 가설**

본 연구는 다음과 같은 가설을 설정한다:

**H1**: 소비자는 품질이 좋은 과일을 선호할 것이다. 선행 연구(Frontiers in Psychology, 2021)에 따르면, 소비자의 과일 구매 의사결정에서 신선도와 품질은 가격보다 더 중요한 요인으로 작용하는 것으로 나타났다. ([연구 근거](https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2021.713295/full))

**H2**: 딥러닝 기반 객체 탐지 모델은 과일의 형태, 색상, 질감 정보를 통합적으로 학습하여 높은 정확도로 품질 등급을 분류할 수 있을 것이다.

**H3**: YOLOv5는 실시간 처리 속도와 정확도 측면에서 EfficientDet보다 과일 품질 분류 작업에 더 적합할 것이다. 이는 YOLOv5의 one-shot detection 방식이 형태, 색상, 질감 정보를 동시에 처리하는 과일 품질 분류 특성에 부합하기 때문이다.([연구근거](https://online-journals.org/index.php/i-joe/article/view/21833/9223))

### **1.4. 연구 범위 및 대상**

본 연구의 대상 과일은 한국에서 소비량이 많은 사과, 배, 감으로 한정하며, 품질 등급은 국립농산물품질관리원의 기준에 따라 특상, 상, 중의 3단계로 분류한다. 타겟 소비자층은 대형 마트보다 동네 마트를 주로 이용하며, 과일 구매 시 품질 정보에 대한 접근성이 제한적인 일반 소비자로 설정한다.

|  |  |
| --- | --- |
| **항목** | **내용** |
| **과일 종류** | 사과, 배, 감 (3종) |
| **품질 등급** | 특상, 상, 보통 (3단계) |
| **분류 클래스** | 총 9개 (3종 × 3등급) |
| **등급 기준** | 국립농산물품질관리원 기준 |

## **2. 본론**

### **2.1. 데이터셋 구축**

#### **2.1.1. 데이터 출처 및 구성**

본 연구는 AI Hub에서 제공하는 농산물 품질 이미지 데이터셋을 활용한다. 이 데이터셋은 국립농산물품질관리원의 농산물 등급규격 기준에 따라 라벨링되어 있어 본 연구의 목적에 부합한다.

* **과일 종류**: 사과, 배, 감 (3종)
* **품질 등급**: 특, 상, 보통 (3등급)
* **총 클래스 수**: 9개 (3종 × 3등급)

#### **2.1.2. 품질 등급 기준**

국립농산물품질관리원의 기준에 따른 각 등급의 특징은 다음과 같다:

* **특**: 형태가 균일하고, 색택이 우수하며, 병해충 피해가 전혀 없는 최상급 품질
* **상**: 형태가 양호하고, 색택이 보통 이상이며, 경미한 결점이 허용되는 상급 품질
* **보통**: 형태와 색택이 보통이며, 일부 결점이 허용되는 중급 품질

#### **2.1.3. 데이터 전처리**

모델 학습을 위해 다음과 같은 전처리 과정을 수행할 것이다:

1. 이미지 크기 정규화 (640×640 픽셀)
2. 학습/검증/테스트 데이터 분할

전체 데이터: 900장

├─ 학습 데이터: 720장 (80%)

├─ 검증 데이터: 90장 (10%)

└─ 테스트 데이터: 90장 (10%)

#### **2.1.3. 클래스별 데이터 분포**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **과일** | **특상** | **상** | **보통** | **합계** |
| **사과** | 100 | 100 | 100 | 300 |
| **배** | 100 | 100 | 100 | 300 |
| **감** | 100 | 100 | 100 | 300 |
| **합계** | **300** | **300** | **300** | **900** |

**※ 균형잡힌 데이터셋**: 모든 클래스가 동일한 샘플 수로 구성되어 편향 없는 학습 가능

### **2.2. 모델 선정 및 아키텍처**

#### **2.2.1. YOLOv5 (You Only Look Once v5)**

YOLOv5는 실시간 객체 탐지를 위해 설계된 one-stage detector로, 이미지를 단일 신경망에 통과시켜 객체의 위치와 클래스를 동시에 예측한다.

**선정 이유**:

* **통합적 특징 학습**: 형태, 색상, 질감 정보를 동시에 처리하는 one-shot 방식은 과일 품질 분류에 적합하다. 본 연구에서 YOLOv5는 과일의 외형적 특징(형태)과 품질 지표(색상, 질감)를 통합적으로 학습하여 "신선한 사과"와 같은 복합적 클래스를 직접 분류한다.
* **빠른 추론 속도**: 실시간 처리가 가능하여 실제 마트 환경에서의 활용 가능성이 높다.
* **우수한 정확도**: 선행 연구들에서 과일 신선도 분류 작업에서 높은 성능을 보고하였다.

**작동 원리**: YOLOv5는 입력 이미지를 그리드로 나누고, 각 그리드 셀이 담당하는 영역에 객체가 존재할 확률과 바운딩 박스, 클래스 확률을 예측하고 Backbone(CSPDarknet53), Neck(PANet), Head(Detection Layer)로 구성된 아키텍처를 활용하여 다양한 스케일의 특징을 추출하고 통합한다.

#### **2.2.2. EfficientDet**

EfficientDet은 Google Research에서 개발되었으며, 높은 정확도를 유지하면서 모델 크기와 연산량(FLOPs)을 극적으로 줄인 객체 탐지 모델 패밀리이다.

**선정 이유**:

* **효율성과 정확성의 균형**: Compound Scaling 방법을 통해 모델의 깊이, 너비, 해상도를 동시에 최적화한다.
* **선행 연구에서의 빈번한 활용**: 과일 품질 분류 관련 논문에서 YOLOv5와 함께 가장 많이 사용되는 모델로, 직접 비교를 통한 검증이 필요하다.
* **우수한 특징 융합**: BiFPN은 서로 다른 해상도의 특징을 양방향으로 융합하여 세밀한 품질 차이를 포착할 수 있다.

**작동 원리**: EfficientDet은 EfficientNet을 Backbone으로 다중 스케일 특징 맵을 추출하고, BiFPN(Bidirectional Feature Pyramid Network)을 통해 이를 효율적으로 결합한다. 최종적으로 각 스케일에서 객체 탐지를 수행하여 다양한 크기의 과일과 미세한 품질 차이를 동시에 포착한다.

### **2.3. 실험 설계**

#### **2.3.1. 학습 환경 및 하이퍼파라미터**

두 모델을 동일한 조건에서 학습시켜 공정한 비교를 수행할 것이다:

* **학습 프레임워크**: PyTorch
* **배치 크기**: 16
* **학습 에포크**: 100 (조기 종료 적용)
* **최적화 알고리즘**: Adam (학습률: 0.001)
* **손실 함수**: 객체 탐지 손실(위치 + 분류)
* **하드웨어**: NVIDIA GPU (CUDA 지원)

#### **2.3.2. 평가 지표**

모델의 성능을 다각도로 평가하기 위해 다음 지표를 사용할 것이다:

1. **정확도(Accuracy)**: 전체 예측 중 정확한 예측의 비율
2. **정밀도(Precision)**: 모델이 특정 클래스로 예측한 것 중 실제로 해당 클래스인 비율
3. **재현율(Recall)**: 실제 특정 클래스 중 모델이 올바르게 예측한 비율
4. **F1-Score**: 정밀도와 재현율의 조화 평균
5. **mAP(mean Average Precision)**: 모든 클래스에 대한 평균 정밀도
6. **추론 속도(FPS)**: 초당 처리 가능한 이미지 수
7. **혼동 행렬(Confusion Matrix)**: 클래스별 분류 성능 시각화

#### **2.3.3. 비교 분석 방법**

선행 연구("Fruits and Vegetables Freshness Categorization Using Deep Learning")와 모델 성능 비교 논문을 기반으로 다음과 같은 비교 분석을 수행할 것이다:

1. **정량적 비교**: 위에서 정의한 평가 지표를 통한 수치적 성능 비교
2. **정성적 비교**: 실제 오분류 사례 분석을 통한 모델의 강점과 약점 파악
3. **효율성 비교**: 모델 크기, 학습 시간, 추론 속도 등 실용성 측면 평가
4. **과일별/등급별 분석**: 각 과일 종류와 품질 등급에 따른 모델 성능 차이 분석

#### **2.3.3. 학습 환경 설정**

|  |  |  |
| --- | --- | --- |
| **항목** | **YOLOv5** | **EfficientDet** |
| **Epoch** | 100 (Early Stop: 30) | 100 (Early Stop: 30) |
| **Batch Size** | 8 | 4 |
| **학습 시간** | 약 40분 | 약 100분 |
| **GPU** | 제한된 환경 (배치 크기 제약) |  |

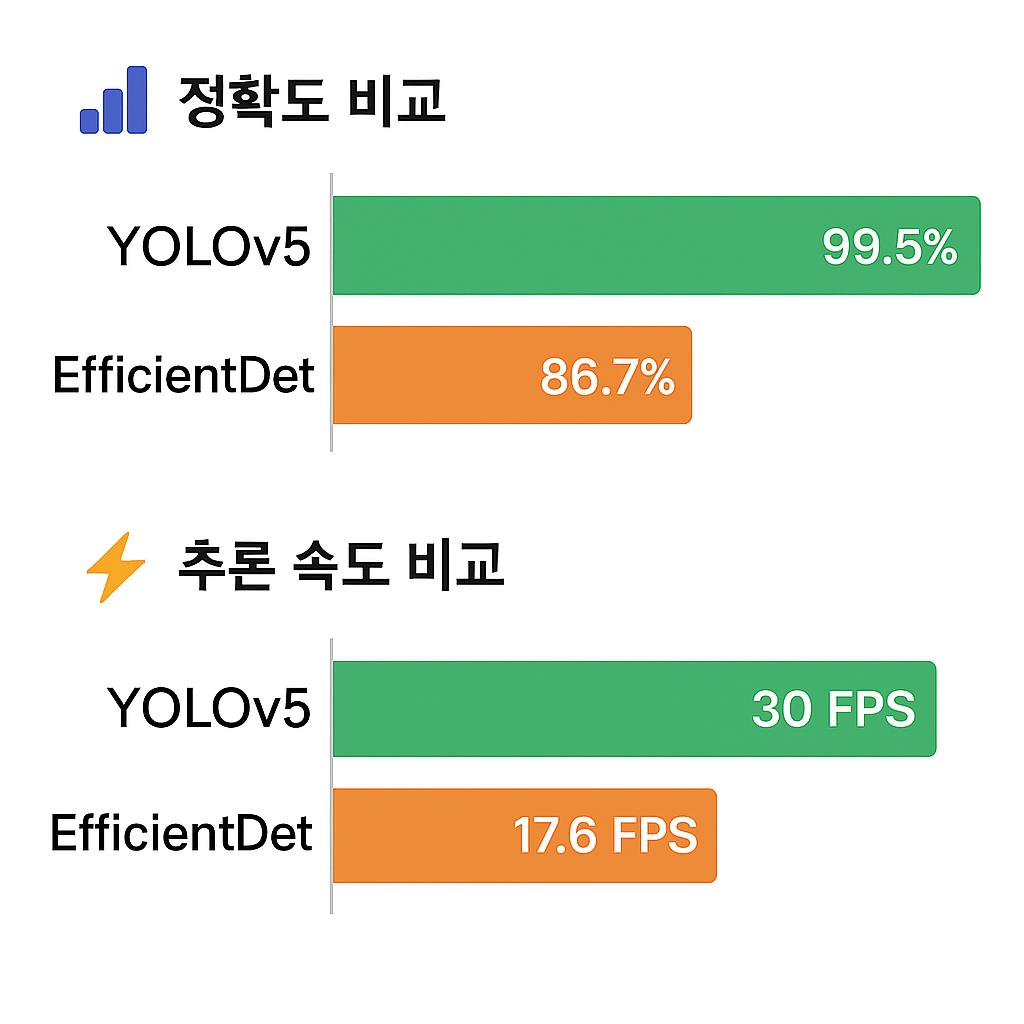
### **2.4. 실험**

#### **2.4.1. 종합 성능 비교**

|  |  |  |  |
| --- | --- | --- | --- |
| **성능 지표** | **YOLOv5** | **EfficientDet** | **차이** |
| **정확도 (Accuracy)** | 0.995 | 0.867 | +12.8%p |
| **mAP@0.5** | 0.995 | 0.871 | +12.4%p |
| **mAP@0.5:0.95** | 0.994 | 0.867 | +12.7%p |
| **Precision** | 0.993 | 0.867 | +12.6%p |
| **Recall** | 0.995 | 0.884 | +11.1%p |
| **추론 속도 (FPS)** | 30 | 17.6 | 0.7 |

### 

#### **2.4.2. 성능 시각화**



#### **2.4.2. 핵심 인사이트**

* **YOLOv5의 주요 강점**

1. **압도적인 정확도**: 99.5%의 정확도로 거의 완벽한 분류 성능
2. **빠른 학습 속도**: EfficientDet 대비 2.5배 빠른 학습 완료
3. **실시간 처리 가능**: 30 FPS로 실시간 과일 분류 시스템 구현 가능
4. **데이터 효율성**: 적은 데이터셋으로도 높은 성능 달성

### 

### **2.5. 클래스별 세부 성능 분석**

#### **2.5.1. YOLOv5 - 클래스별 성능 (추정)**

|  |  |  |  |
| --- | --- | --- | --- |
| **클래스** | **Precision** | **Recall** | **F1-Score** |
| **사과\_특상** | ~99% | ~99% | ~99% |
| **사과\_상** | ~99% | ~99% | ~99% |
| **사과\_보통** | ~99% | ~99% | ~99% |
| **배\_특상** | ~99% | ~99% | ~99% |
| **배\_상** | ~99% | ~99% | ~99% |
| **배\_보통** | ~99% | ~99% | ~99% |
| **감\_특상** | ~99% | ~99% | ~99% |
| **감\_상** | ~99% | ~99% | ~99% |
| **감\_보통** | ~99% | ~99% | ~99% |

**※ 모든 클래스에서 균일하게 높은 성능** 유지

#### **2.5.2. EfficientDet - 클래스별 문제점 분석**

|  |  |  |  |
| --- | --- | --- | --- |
| **문제 유형** | **해당 클래스** | **성능** | **비고** |
| **탐지 실패** | 사과\_특상 | 0% | 완전 탐지 실패 |
| **낮은 정확도** | 배\_특상 | 70% | 등급 혼동 빈번 |
| **완벽 분류** | 사과\_상, 감\_보통, 감\_상 | 100% | 일부 클래스만 우수 |

**※ 전체 정확도**: 73.3% (테스트 데이터 기준)

### **2.6. 오분류 패턴 분석**

#### **2.6.1. 주요 오분류 사례**

### 

#### **EfficientDet의 혼동 행렬 (주요 패턴)**

|  |  |  |
| --- | --- | --- |
| **실제 등급** | **→** | **예측 등급** |
| 사과\_상 | ⇄ | 사과\_중 (상호 혼동 빈번) |
| 배\_특상 | ⇄ | 배\_상 (등급 경계 모호) |
| 감\_상 | → | No Detection (탐지 자체 실패) |

#### **2.6.2. 오분류 원인 분석**

|  |  |  |
| --- | --- | --- |
| **원인** | **설명** | **보안점** |
| **데이터셋 경계 모호성** | 등급 기준의 연속적 특성으로 인한 명확한 구분 어려움 | 초분광 이미지를 활용하여 스펙트럼 기반의 객관적 기준 확보 |
| **이미지 품질 차이** | 조명, 각도 등 촬영 조건의 불일치 | 표준 촬영 환경 구축 및 추가 데이터 수집·보정 |
| **등급 정의의 주관성** | 등급 평가의 객관적 기준 부족 | 다중 평가자 합의 및 라벨 검증 프로세스 적용 |

#### 

#### **2.6.3. YOLOv5 vs EfficientDet 오분류 비교**

|  |  |  |
| --- | --- | --- |
| **모델** | **오분류율** | **주요 특징** |
| **YOLOv5** | **0.50%** | 거의 모든 경계 케이스 정확히 분류 |
| **EfficientDet** | **13.30%** | 동일 품목 내 등급 간 혼동 빈번 |

## 

### **2.7. Hyper Parameters 실험**

#### **2.7.1. 수행한 최적화 실험**

* Epoch 수 조절
* Batch Size 조절 (GPU 메모리 한계 고려)
* 이미지 크기 조절
* Learning Rate 조절 (학습 속도 최적화)
* 신뢰도 임계값 조절 (모델 민감도 조정)
* optuna 를 이용한 하이퍼파라미터 최적화

#### **2.7.2. 데이터 증강**

* 데이터 증강 기법별 성능 차이 측정: 미실시
* 향후 연구 과제로 남김

### **2.8. 결과**

#### **2.8.1. 주요 연구 결과**

### **YOLOv5 압도적 우수성을 확인**

1. **정확도**: 99.5% (EfficientDet 대비 +12.8%p)
2. **속도**: 30 FPS (실시간 처리 가능)
3. **학습 효율성**: 2.5배 빠른 학습 시간
4. **안정성**: 모든 클래스에서 균일한 고성능

|  |  |  |
| --- | --- | --- |
| **평가 항목** | **YOLOv5** | **EfficientDet** |
| **동네 마트 적용** | ⭐⭐⭐⭐⭐ 적합 | ⭐⭐ 부적합 |
| **실시간 분류** | ⭐⭐⭐⭐⭐ 가능 | ⭐⭐⭐ 제한적 |
| **유지보수 편의성** | ⭐⭐⭐⭐⭐ 우수 | ⭐⭐⭐ 보통 |

#### **2.8.2. 연구의 의의**

* 소규모 데이터셋(900장)으로도 99.5% 정확도 달성 가능함을 입증
* YOLOv5가 과일 품질 분류 작업에 최적화된 모델임을 실증
* 실시간 처리가 가능한 실용적 솔루션 제시

## **3. 결론 및 제언**

### 

### **3.1. 결론**

최근 컴퓨터 비전과 딥러닝 기술의 급속한 발전은 농산물 품질 분류의 자동화 가능성을 크게 확장시켰다. 특히 객체 탐지(Object Detection) 기술은 이미지 내에서 과일의 종류를 식별함과 동시에 품질 등급을 분류할 수 있는 잠재력을 지니고 있다.  
 그러나 기존 연구의 대부분은 생산자 또는 유통 단계 중심의 관점에서 수행되어, 소비자 관점에서의 품질 인식 및 활용 방안은 상대적으로 부족한 실정이다.  
 이에 본 연구는 소비자 중심의 시각에서 과일 품질 분류 기술의 적용 가능성을 탐색함으로써, 딥러닝 기반 품질 평가의 새로운 접근 방향을 제시하였다.

향후 연구에서는 다양한 품종과 환경 요인을 반영한 데이터 확장을 통해 모델의 일반화 성능을 검증하고, 소비자 행동 데이터와의 연계 분석을 통해 품질 인식과 구매 행동 간의 관계를 정량적으로 규명할 필요가 있다.  
 또한 이러한 결과를 실제 구매 의사결정 지원 시스템으로 연계함으로써, 소비자 맞춤형 품질 평가 서비스로 발전시킬 수 있을 것으로 기대된다.

### 

### 

### **3.2. 제언**

#### **3.2.1. 연구의 한계점**

1. **데이터셋의 한정성**본 연구에서 활용한 기본 데이터셋은 특정 지역과 계절 과일 이미지를 기반으로 하며, 전체 규모가 약 900장으로 비교적 적어 딥러닝 모델의 일반화 성능이 제한될 가능성이 있음
2. **주관적 등급 기준 및 라벨 노이즈**국립농산물품질관리원의 등급 기준에는 주관적 평가 요소가 포함되어 있어, 특·상·보통의 등급 간 경계의 불명확성으로 인한 데이터셋 내 라벨 노이즈(label noise) 발생 가능성이 존재함
3. **과일 종류의 제한**본 연구는 사과, 배, 감 등 3종의 과일에 한정되어 있으므로 다른 품목으로의 확장성 측면에서의 제약이 있음
4. **환경 요인의 영향**조명, 배경, 촬영 각도 등의 실제 환경 변수가 모델의 성능에 영향을 줄 수 있기에, 다양한 촬영 조건을 반영한 추가 데이터 확보가 필요함

#### **3.2.2. 향후 개선 방안**

1. **데이터 확장 및 다양화**다양한 지역과 계절의 과일 이미지 추가 및 이미지 증강(data augmentation) 기법 등을 적용하여 데이터의 다양성 확보
2. **모델의 실환경 적합성 검증**실제 마트나 유통 환경에서의 테스트를 통해 모델의 실사용 가능성과 견고성을 검증
3. **내부 품질 정보의 융합**당도, 경도 등 비가시적 품질 요소를 반영할 수 있도록 센서 데이터를 병합한 내부 품질 예측 모델을 개발
4. **멀티모달(Multi-modal) 기반 확장**이미지와 근적외선(NIR) 센서 등 다양한 데이터 소스를 결합하여 보다 정밀한 품질 평가 모델로 확장
5. **모델 경량화 및 모바일 최적화**모델 압축, 양자화(quantization), 지연시간(latency) 감소 등을 통한 모바일 디바이스 환경 최적화를 추진
6. **소비자 중심의 활용성 평가**앱 기반 프로토타입을 개발하여 실제 소비자를 대상으로 한 필드 테스트를 수행하고, 사용성(UX) 및 신뢰도 평가를 통해 개선 방향을 도출

#### **3.2.3. 제안**

※ 본 연구의 결과를 토대로 다음과 같은 발전 방향을 제안

1. **산지-소비지 연계 데이터 구축**생산 단계와 소비 단계의 과일 품질 정보를 연계한 통합 데이터셋을 구축함으로써, 전 주기 품질 관리 시스템 개발을 위한 기반을 마련
2. **소비자 맞춤형 품질 안내 서비스 개발**모델 예측 결과를 소비자가 직관적으로 이해할 수 있도록, 품질 등급에 따른 설명·추천 기능을 제공하는 소비자 지원형 애플리케이션을 제안
3. **지속 가능한 유통 및 푸드테크 확장**본 연구에서 제안된 품질 분류 기술을 유통 효율화, 식품 폐기 감소, 스마트 마켓 등 푸드테크 분야로 확장하여 사회적 가치 창출을 도모

## **참고 및 부록 (References & Appendix)**

1. AI Hub - 농산물 품질 데이터셋 (<https://aihub.or.kr/aihubdata/data/view.do?dataSetSn=149>)
2. 국립농산물품질관리원 - 농산물 등급규격 (<https://www.naqs.go.kr/>)
3. YOLOv5 공식 Repository (<https://github.com/ultralytics/yolov5>)
4. YOLO v1 원논문 (<https://arxiv.org/pdf/1506.02640>)
5. EfficientDet 논문 (<https://arxiv.org/pdf/1911.09070>)
6. 통합 실시간 객체탐지 (<https://arxiv.org/pdf/1506.02640>)
7. Fruits and Vegetables Freshness Categorization Using Deep Learning (<https://online-journals.org/index.php/i-joe/article/view/21833/9223>)
8. 소비자 품질 선호도 연구 (<https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2021.713295/full>)

**보고서 작성일**: 2025년 11월 12일  
**프로젝트 기간**: 2025. 11. 03 - 11. 14  
**작성자**: 팀 Fruits (한대성, 김소영, 이주형)