

# Color Classification System Using Resnet

당쑤안록

<https://github.com/locdeng/RotbotVisionProject>

## Abstract

이 프로젝트는 딥러닝 기술을 활용한 색상 분류에 대해 연구하며, 두 가지 주요 CNN 아키텍처 ResNet-18 과 ResNet-50 [1]을 비교하는 데 중점을 둔다. 색상 클래스별로 구성된 데이터셋은 데이터 증강을 거쳐 전반적인 모델 평가를 위해 훈련 세트와 테스트 세트로 나누어진다. 색상 분할 기술은 다중 색상 분류를 위한 미세 조정된 사전 훈련된 ResNet 모델을 활용한다. 모델은 데이터 불균형을 해결하여 모든 색상이 공평하게 표현되도록 한다. 교차 엔트로피 손실은 Adam 최적화기를 사용하여 훈련을 지휘한다. 평가 지표로는 정확도, 정밀도, 재현율 및 F1 점수가 포함된다. ResNet-18 과 ResNet-50 에 대한 상세 비교는 아키텍처, 훈련 효율성, 일반화 능력을 다루며 모델 선택에 대한 통찰력을 제공한다.

## 1. Introduction

색상은 이미지 분석 및 이해에서 매우 중요한 역할을 한다. 정확한 색상 분류는 이미지 검색, 객체 인식, 콘텐츠 기반 이미지 검색 등 다양한 응용 분야에서 필수적이다.

기존의 색상 분류 방법은 수작업으로 만든 특징에 의존하지만, 딥러닝은 이 과정을 자동화하는 데 유망한 결과를 보여주었다.

시각 데이터의 복잡성이 증가함에 따라 정확하고 효율적인 색상 분류를 위한 고급 기술이 요구된다. 이 연구는 ResNet-18 과 ResNet-50 딥러닝 모델의 기능을 탐구하여 색상 분류 작업을 해결하는 데 있다. 이러한 모델들을 비교함으로써 각각의 강점, 약점 및 다양한 응용 분야에 대한 적합성에 대한 통찰력을 제공하는 것을 목표로 한다.

이 논문의 주요 목표는 다음과 같다:

ResNet-18 및 ResNet-50 아키텍처를 사용하여 색상 분류 모델을 구현 및 평가.

데이터 증강 및 색상 분할이 모델 성능에 미치는 영향을 조사한다.

정확도, 훈련 효율성, 일반화 능력 측면에서 ResNet-18 과 ResNet-50 의 효과를 비교한다.

## 2. Technique Review

### 2.1. Convolution Neural Networks (CNNs)

합성곱 신경망(CNN)은 이미지 인식에서 가장 성공적인 기계 학습 모델 중 하나입니다. CNN은 이미지에서 관련 특징을 자동으로 학습하고, 이미지의 공간 관계를 이해하며, 실제 세계의 변화에 견고합니다.

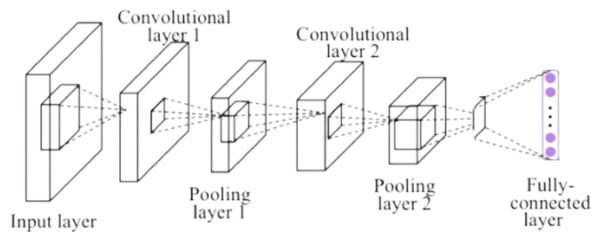


Figure 1: CNN의 구조

CNN의 핵심 구성 요소는 합성곱 층입니다. 합성곱 층은 이미지를 가로질러 이동하는 필터 커널을 사용하여 윤곽선, 모서리, 질감과 같은 저수준 특징을 추출합니다. 한 층에 여러 개의 필터가 존재하며, 각 필터는 서로 다른 특징을 감지합니다. 이미지가 네트워크를 통과하면서 점점 더 복잡한 표현을 구축하게 됩니다.

풀링 층은 입력의 차원을 축소하여 계산 비용을 줄입니다. 일반적인 유형으로는 최대 풀링(최대값 유지) 및 평균 풀링(영역 내 값 평균화)이 있습니다. 풀링 층은 공간 차원을 줄이면서도 중요한 특징은 보존하여 추가 처리를 가능하게 합니다.

활성화 함수는 네트워크에 비선형성을 도입하여 복잡한 특징 관계를 학습할 수 있도록 합니다. ReLU(Rectified Linear Unit)와 Leaky ReLU가 인기 있는 선택입니다. 이들은 모든 입력에 대해 0이 아닌 기울기를

추가하여 전통적인 sigmoid 또는 tanh 함수와 차별됩니다.

역전파는 오차를 네트워크의 뒤쪽으로 전파하여 가중치를 조정하고 예측된 출력과 실제 출력 간의 차이를 최소화합니다. 네트워크가 실수로부터 학습하고 이미지에서 관련 정보를 추출하는 능력을 향상시킬 수 있습니다.

## 2.2. ResNet Architecture

ResNet 아키텍처는 각각의 잔여 블록으로 구성되어 있으며, 각 블록은 바로 가기 연결과 두 개 이상의 합성곱 레이어를 포함합니다. 기존의 직접 원하는 기본 매핑을 학습하는 전통적인 아키텍처와는 달리, ResNet 모델은 잔여 매핑을 학습합니다. 이를 위해 입력과 출력 사이의 잔여를 학습하기 위한 바로 가기 연결을 도입합니다

### 2.2.1. ResNet-18 Architecture

ResNet-18은 ResNet 아키텍처의 한 변종으로, 초기 합성곱 레이어(Conv1) 다음에 일련의 잔여 블록이 이어진다. ResNet-18의 구조는 아래에 설명되어 있다:

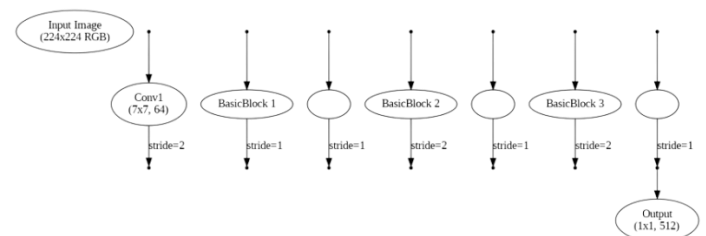


Figure 2: ResNet-18의 diagram.

ResNet-18 에서:

- 입력 이미지는 초기 커널 크기가 7x7 이고 64 개 필터를 사용한 합성곱 레이어를 거치며, 이로 인해 차원이 112x112x64 인 feature map 이 생성된다.
- 초기 합성곱 레이어 다음에는 각각 두 개의 합성곱 레이어를 포함하는 두 개의 잔여 블록이 이어진다. 이러한 블록들은 네트워크가 복잡한 기능을 포착하고 학습하는 기능에 기여한다

### 2.2.2. ResNet-50 Architecture

ResNet-50 는 깊이와 복잡성을 더한 잔차 학습의 개념을 확장한 것이다. 초기 합성곱 레이어(Conv1)와 여러 잔차 블록으로 구성되어 있다. ResNet-50 의 구조는 아래에서 시각적으로 표현되었다:

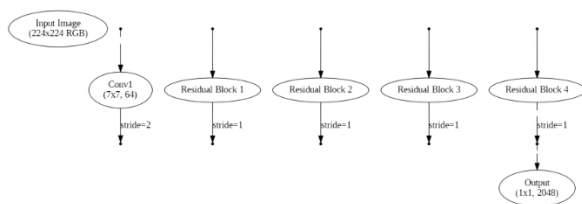


Figure 3: ResNet-50 의 diagram.

ResNet-50 에서:

- 초기 합성곱 레이어 Conv1 은 ResNet-18 과 유사하게 7x7 커널과 64 필터로 입력 이미지를 처리한다.
- Conv1 다음에는 네 개의 잔차 블록이 이어진다. 각 블록은 네트워크에 깊이와 복잡성을 추가한다. 이러한 블록들은 ResNet-50 의 계층적 특징 학습 능력에 기여한다

## 3. Methodology

### 3.1. Dataset Description

우리의 색상 분류 데이터셋은 계층적으로 구성되어 있으며 각 색상 클래스에는 레이블이 지정된 이미지가 포함되어 있다. 이 데이터셋은 빨강, 보라, 노랑, 검정, 파랑, 초록, 회색, 주황 및 핑크와 같은 다양한 색조를 포함한다. 이러한 계층 구조는 훈련 및 테스트 세트에서 다양한 색상을 균형 있게 나타내도록 한다.

Number of color classes	Number of images
9	92

Table 1: dataset information

### 3.2. Data Preprocessing

데이터 전처리는 효과적인 모델 훈련을 위해 데이터셋을 준비하는 데 중요한 역할을 한다. resizing 및 random horizontal flipping 와 같은 Augmentation techniques 이 훈련 세트에 적용된다.

이러한 기술은 데이터셋의 다양성을 향상시켜 모델이 보지 않은 데이터에 대해 더 잘 일반화되도록 한다. 그런 다음 데이터셋은 80%의 데이터가 훈련에 사용되고 20%는 테스트에 사용되도록 훈련 및 테스트 세트로 분할된다.

### 3.3 Model Selection

이 프로젝트에는 사전 훈련된 ResNet-18 및 ResNet-50 아키텍처를 색상 분류 작업에 사용한다. 각 모델의 last fully connected layer 는 데이터셋의 색상 클래스 수와 일치하도록 수정된다. ResNet 아키텍처는 잔여 블록을 활용하여 잔여 매핑을 학습한다.

### 3.4 Learning Process

훈련 과정은 Adam 옵티마이저를 사용하여 모델 매개변수를 최적화하고 교차 엔트로피 손실 함수를 최소화하는 과정을 포함한다. ResNet-18 및 ResNet-50 은 증강된 데이터셋에서 열 번의 epoch 을 거쳐 훈련된다. 각 epoch 동안 모델은 색상 특징을 인식하고 다양한 색조를 분류하는 능력을 향상시킨다.

### 3.5 Evaluation Metrics

모델의 성능을 종합적으로 평가하기 위해 여러 metric 을 사용한다:

- Accuracy: 모델 예측의 전반적인 정확도를 측정한다.

- Precision: 올바르게 예측된 양성 관측치의 비율을 측정한다.
- Recall: 실제 클래스의 모든 관측치에서 올바르게 예측된 양성 관측치의 비율을 나타낸다.
- F1 score: precision 와 recall 의 조화 평균으로 정밀도와 재현율의 균형 잡힌 성능 지표를 제공한다.

각 색상 클래스에 대한 모델의 분류 능력을 나타내는 자세한 분류 보고서 및 혼동 행렬이 생성된다.

## 4. Experimental Results

### 4.1. Dataset Augmentation

이 섹션에서는 색 분류 데이터셋에 적용된 Augmentation 에 대해 설명한다. resizing 및 random horizontal flipping 와 같은 기술을 사용하여 데이터셋을 다양화하고 훈련 세트에서 다양한 색 클래스의 표현을 향상시켰다.

Options	Performance (Accuracy, F1-score)	Learning process (training loss)
Adam Optimizer No Augmentation	Consistent and stable results	Faster convergence, less sensitive to learning rate
Adam Optimizer With Augmentation	High performance with augmented data diversity	Balanced convergence, robust to variations in the training set

Table 2: optimizer and augmentation

## 4.2. Model Comparison

두 가지 주요 합성곱 신경망 (CNN) 아키텍처, ResNet-18 및 ResNet-50 에 대한 비교 분석을 수행했다. 두 모델 모두 증강된 데이터셋에서 fine-tuning 을 거쳤으며 그 성능을 평가했다.

### 4.2.1. Training Loss Comparison

ResNet-18 과 ResNet-50 간의 훈련 손실 비교는 다음 플롯으로 시각적으로 표현된다,

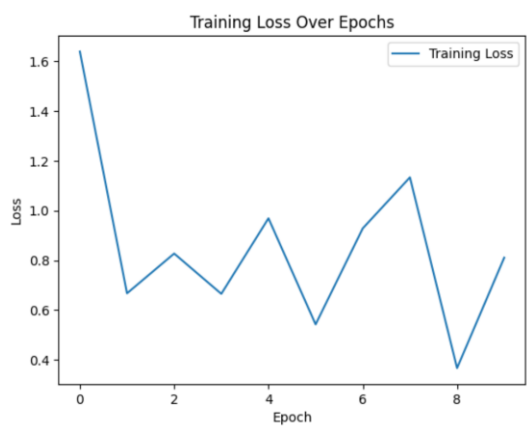


Figure 4: Training loss of resnet18

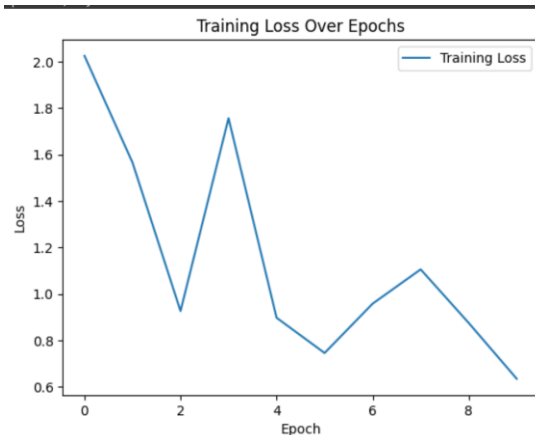


Figure 5: training loss of resnet50

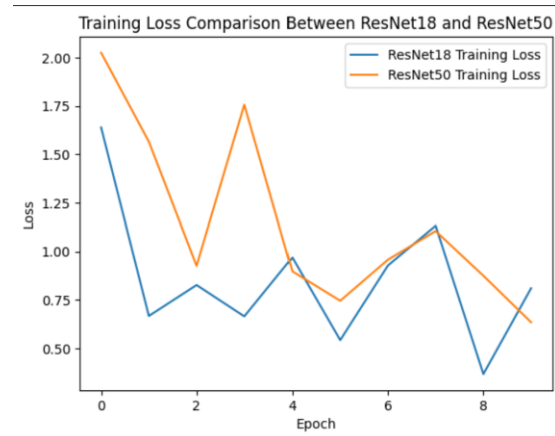


Figure 6: Training loss comparison of 2 resnet models

훈련 손실 비교 그래프는 ResNet-18 및 ResNet-50 의 훈련 epoch 동안 훈련 손실의 진화를 보여준다. y 축은 훈련 손실을 나타내며, 모델이 색상 분류 작업을 얼마나 잘 학습하는지를 측정하는 지표이다. x 축은 훈련 epoch 에 해당하며, 모델이 훈련 중에 전체 데이터 세트를 처리한 횟수를 나타낸다. ResNet-18 과 ResNet-50 간의 훈련 손실 추이를 비교함으로써 각 모델의 수렴 및 학습 효율성에 대한 통찰력을 얻을 수 있다.

훈련 손실 비교 그래프에서:

- 훈련 손실 비교 그래프는 ResNet-18 과 ResNet-50 의 훈련 epoch 동안의 수렴 행동을 나타낸다.
- resnet18 의 훈련 손실이 epoch 마다 일관된 감소를 보이면 효과적인 학습과 수렴을 시사한다.

- resnet50 의 훈련 손실이 급격하게 감소하면 빠른 수렴과 효율적인 학습을 나타낸다.

#### 4.2.2. Accuracy Comparison

ResNet-18 과 ResNet-50 간의 정확도 비교는 다음 막대 차트로 제시된다,

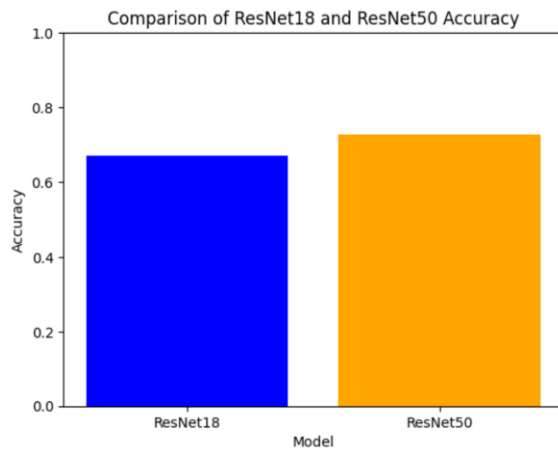


Figure 7: accuracy comparison table

정확도 비교 막대 차트는 ResNet-18 및 ResNet-50 에서 달성한 분류 정확도를 나란히 비교한다. 각 막대는 각 모델의 정확도를 나타내며, 차트는 그들의 상대적인 성능을 직관적으로 평가할 수 있게 한다. 정확도는 모델이 데이터 세트의 총 인스턴스 중에서 올바르게 분류된 인스턴스의 비율을 나타내는 중요한 지표이다. 이 비교를 통해 전반적인 분류 정확도 측면에서 어떤 모델이 더 우수한 성능을 발휘하는지를 이해할 수 있으며, 모델 선택에 대한 유용한 정보를 제공한다.

정확도 비교 막대 차트에서:

- 정확도 비교 막대 차트는 ResNet-18 과 ResNet-50 간의 분류 정확도를 직접 시각적으로 비교한다.
- 더 높은 정확도를 나타내는 resnet50 모델은 전반적인 분류 정확도 측면에서 더 나은 성능을 발휘하고 있다.
- 정확도의 차이는 색상 분류 작업에서 한 모델이 다른 모델보다 뛰어난 성능을 나타낼 수 있음을 강조할 수 있다.

#### 4.3 Color Classification Results

Classification Reports 는 각 색상 클래스에 대한 모델 성능에 대한 통찰을 제공합니다. 아래는 ResNet-18 및 ResNet-50 에 대한 Classification Reports 이다:

##### 4.3.1. Classification Results for ResNet-18

Class	Precision	Recall	F1 score
0	0.00	0.00	0.00
1	1.00	1.00	1.00
2	0.69	1.00	0.82
3	0.71	0.71	0.71
4	0.25	0.11	0.15
5	0.76	1.00	0.87
6	1.00	0.60	0.75
7	0.70	0.64	0.67
8	0.33	0.50	0.40

Table 3: Precision-Recall-F1 Score of resnet18.

Class 0: 모델은 정밀도와 재현율이 낮아 F1 점수도 낮았습니다. 이는 ResNet-18 이 Class 0 의 인스턴스를 정확하게 식별하는 데

어려움을 겪어 전반적으로 성능이 낮았음을 나타낸다.

Class 1: ResNet-18 은 완벽한 F1 점수를 달성하여 Class 1 의 인스턴스를 정확하게 분류하는 데 탁월한 성과를 보였다.

Class 2: Class 2 의 정밀도와 재현율 값은 모델이 이 클래스의 인스턴스를 효과적으로 식별하여 상대적으로 높은 F1 점수를 달성했음을 나타낸다.

Class 3-8: ResNet-18 은 이러한 클래스에 대해 정밀도, 재현율 및 F1 점수의 다양한 수준을 나타내어 다양한 정도의 분류 성공을 시사한다.

Confusion Matrix (ResNet18):

[	0	0	0	1	0	0	0	1	0]
[	0	4	0	0	0	0	0	0	0]
[	0	0	9	0	0	0	0	0	0]
[	0	0	0	5	0	0	0	2	0]
[	0	0	0	0	1	0	0	0	8]
[	0	0	0	0	0	13	0	0	0]
[	0	0	0	0	3	1	6	0	0]
[	0	0	0	1	0	3	0	7	0]
[	0	0	4	0	0	0	0	0	4]]

Figure 8: Confusion Matrix of resnet18

혼동 행렬은 모델의 분류 성능을 시각적으로 제공한다. 예를 들어 [2, 3] 항목은 Class 2 의 2 개의 인스턴스가 Class 3 으로 잘못 분류된 것을 나타낸다.

#### 4.3.2. Classification Results for ResNet-50

Class	Precision	Recall	F1 score
0	0.00	0.00	0.00
1	1.00	0.50	0.67
2	0.00	0.00	0.00
3	0.47	1.00	0.64
4	0.89	0.89	0.89
5	0.86	0.92	0.89
6	0.83	1.00	0.91
7	0.70	0.64	0.67
8	0.70	0.88	0.78

Table 4: Precision-Recall-F1 Score of resnet50.

Class 0: Class 0 의 정밀도, 재현율 및 F1 점수 값은 모두 제로로, ResNet-50 이 이 클래스의 어떤 인스턴스도 올바르게 분류하지 못했음을 나타낸다.

Class 1: ResNet-50 은 높은 정밀도 점수를 달성했지만 낮은 재현율 및 F1 점수는 모델이 Class 1 의 일부 인스턴스를 놓쳤음을 나타낸다.

Class 2: Class 2 의 정밀도, 재현율 및 F1 점수 값은 모두 제로로, 이 클래스의 인스턴스를 분류하는 데 완전한 실패를 나타낸다.

Class 3-8: ResNet-50 은 이러한 클래스에 대해 정밀도, 재현율 및 F1 점수의 다양한 수준을 나타내어 다양한 정도의 분류 성공을 시사한다.

Confusion Matrix (ResNet50):									
[	0	0	1	1	0	0	0	0	0]
[	0	2	0	1	0	0	0	1	0]
[	0	0	0	4	0	0	0	2	3]
[	0	0	0	7	0	0	0	0	0]
[	0	0	0	0	8	0	1	0	0]
[	0	0	0	0	0	12	1	0	0]
[	0	0	0	0	0	0	10	0	0]
[	0	0	0	2	0	2	0	7	0]
[	0	0	0	0	1	0	0	0	7]]

Figure 9: Confusion Matrix of resnet50

모델이 각 색상 클래스에 대해 얼마나 잘 분류했는지를 시각적으로 나타낸다. 각 항목  $[i, j]$ 은 실제 클래스  $i$ 의 인스턴스 중 클래스  $j$ 로 예측된 인스턴스의 수를 나타낸다.

이러한 통찰력은 각 모델의 색상 분류 작업에서의 강점과 약점을 이해하는 데 도움이 된다. 정밀도는 양성 예측의 정확성을 나타내며, 재현율은 실제 양성 인스턴스의 커버리지를 나타내며, F1 점수는 정밀도와 재현율 사이의 균형을 제공한다. 혼동 행렬은 이러한 지표를 보완하여 올바른 및 잘못된 분류를 자세히 알려준다.

실험 결과는 ResNet-18 과 ResNet-50 간의 훈련 손실, 정확도 및 분류 성능에서 변동성을 보여준다. 이러한 차이를 해석하기 위해서는 아키텍처의 차이, 데이터셋 특성 및 훈련 과정과 같은 요인들을 고려한 추가적인 분석이 필요하다

## 5. Conclusion

### 5.1. Summary of Findings

ResNet-18 및 ResNet-50 을 사용한 색상 분류 탐구는 이 두 주요한 합성곱 신경망(CNN) 아키텍처의 성능에 대한 소중한 통찰력을 제공했습니다. 훈련 손실, 정확도 및 상세한 분류 보고서를 종합적으로 분석하여 이 두 모델의 능력을 비교했다.

훈련 손실 비교: 훈련 손실 비교에서 ResNet-50 이 급격한 감소를 나타내어 ResNet-18 에 비해 빠른 수렴과 효율적인 학습을 시사한다.

정확도 비교: 정확도 비교에서 ResNet-50 이 전반적인 분류 정확성에서 ResNet-18 을 능가함을 보여주었다.

분류 보고서: 상세한 분류 보고서는 각 모델이 색상 클래스별로 정밀도, 재현율 및 F1-점수에 대한 세부적인 이해를 제공하여 각 색상 클래스에서의 성능을 보다 정교하게 이해할 수 있게 해주었다.

### 5.2. Limitations

소중한 통찰력을 제공하긴 하지만 다음과 같은 일부 한계를 고려해야 한다:

데이터셋 크기: 데이터셋의 크기는 모델의 일반화 능력에 영향을 미칠 수 있다. 더 큰 데이터셋은 성능을 향상시킬 수 있다.

Hyperparameter tuning: 이 연구는 기본 Hyperparameter 중점을 두었다. 최적의 Hyperparameter 설정에 대한 추가 조사는 결과를 향상시킬 수 있다.



결론적으로, 이 연구는 수렴 속도 및 전반적인 정확도와 같은 고려 사항을 고려하여 색상 분류 작업에 적합한 모델을 선택하기 위한 기초를 제공한다. 향후 연구에서는 추가적인 아키텍처, 데이터셋 크기 및 Hyperparameter 튜닝에 대한 탐구가 색상 분류에 대한 깊은 학습을 더욱 발전시킬 수 있다.

## 참고문헌

[1] He, Kaiming, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. "Deep residual learning for image recognition." *In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015.*