

## CNN에서 훈련 및 시험 영상 수에 따른 정확도 분석

공준배<sup>○</sup>, 황태희<sup>\*</sup>, 장민석<sup>\*</sup>, 이연식<sup>\*</sup>

국립군산대학교, 컴퓨터정보통신공학부<sup>○</sup>

국립군산대학교, 컴퓨터정보통신공학부<sup>\*</sup>

e-mail: {rhdwnsqo12, thwang, msjang, yslee}@kunsan.ac.kr<sup>○\*</sup>

## Accuracy Analysis according to the Number of Training and Testing Images on CNN

Junbae Kong<sup>○</sup>, Taehee Hwang<sup>\*</sup>, Minseok Jang<sup>\*</sup>, Yonsik Lee<sup>\*</sup>

School of Computer & Information & Communication Engineering,  
College of Engineering, Kunsan National University<sup>○</sup>

School of Computer & Information & Communication Engineering,  
College of Engineering, Kunsan National University<sup>\*</sup>

### ● 요약 ●

본 논문은 CNN (Convolution Neural Networks)의 첫 번째 컨볼루션층(convolution layer)을 RGB-csb(RGB channel separation block)로 대체하여 입력 영상의 RGB 값을 특징 맵에 적용시켜 정확성을 제고시킬 수 있는 선행연구 결과에 추가적으로, 훈련 및 시험 영상 수에 따른 분석을 통하여 정확도 향상 방법을 제안한다. 제안한 방법은 영상의 개수가 작을수록 각 학습 간의 정확도 편차가 크게 나타나는 불안정성은 있지만 기존 CNN모델에 비하여 정확도 차이가 증가함을 알 수 있다.

**키워드:** CNN(Convolution Neural Networks), 정확도(accuracy), RGB채널(RGB channel)

### I. Introduction

최근 학습 및 훈련을 위한 컴퓨팅 자원의 성능이 향상됨에 따라 CNN에 대한 다양한 훈련 방법 및 구조에 대한 GoogLeNet, ResNet, DenseNet, MobileNet과 같은 연구들이 진행되고 있다[1-4]. 또한 신경망의 과적합과 연산량을 줄이기 위한 부가적인 기술을 통하여 정확도를 향상시키기 위한 연구 결과들[5, 6]과 층(layer)의 깊이를 증가시킬 경우 발생하는 vanishing gradient 문제 해결과 정확도 향상을 위하여 연구들[7]이 제시되었으나, 이러한 기술들은 학습효과나 성능저하 문제로 인하여 CNN의 첫 번째 컨볼루션층에 제안 구조를 적용하지 않는다.

본 연구의 선행 연구로써 첫 번째 컨볼루션층을RGB-csb(RGB channel separation block)로 대체하여 RGB-csb를 입력 영상과 가장 가깝게 위치시켜 첫 번째 층에서 정확도를 높일 수 있는 방법을 제안하였고[8], 본 논문에서는 이를 기반으로 훈련 및 시험 영상 수에 따른 분석을 통하여 RGB-csb 대체 모델의 정확도 성능 향상 방법을 제안하고자 한다.

### II. RGB-channel separation block

컬러 영상은 RGB의 3색 정보가 포함되어 있다. 본 논문에서는 영상의 특징을 강조하기 위해 입력영상의 RGB색 정보를 특징맵에 더하여 사용한다. 입력영상은 RGB 3색 정보와 16개 필터의 3x3 컨볼루션을 통해 특징을 더욱 더 견고하게 나타낸다. CNN을 통과하는 하나의 RGB 영상 X를 224 x 224 크기이며  $X_r, X_g, X_b$ 의 3가지 색 정보를 포함하고 있다고 가정한다. 일반적으로 영상 X가 순수한 컨볼루션 연산만을 수행하는 첫 번째 컨볼루션층을 통과하게 하는 대신에, 첫 번째 컨볼루션층을 RGB-csb로 대체한다. 다음 Fig. 1.은 CL(Convolution Layer), SUM, CR(Channel Reduction)로 구성된 사이클을 R, G, B 값에 따라 3번 진행되는 RGB-csb의 구조를 나타낸다.

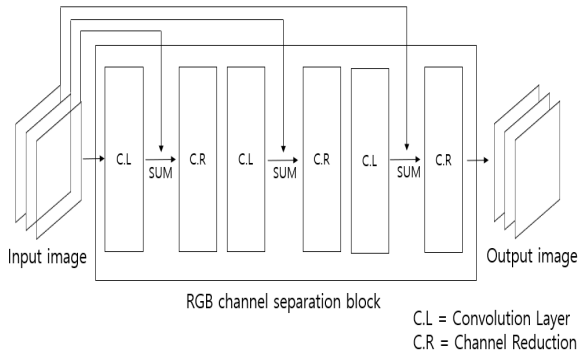


Fig. 1. RGB-channel separation block(RGB-csb) structure

RGB-csb의 구성요소들의 특징을 나타내는 내부구조는 다음 Fig. 2와 같으며, 이를 적용한 학습과정은 다음과 같다[8].

- 1) 영상 X를 작은 수의 필터(본 논문에서는 16개)를 사용한 3 x 3 컨볼루션, BN, ReLU를 차례대로 적용하여 224 x 224 x 16의 특징 맵을 얻는다.
- 2) 특징 맵의 16개 채널들에 대해 SUM연산을 통하여 각각의 Xr을 더한다. (순차적으로 Xg와 Xb에도 동일하게 적용)
- 3) 16채널을 3개의 채널로 줄이기 위하여 CR(Channel Reduction)을 수행한다.

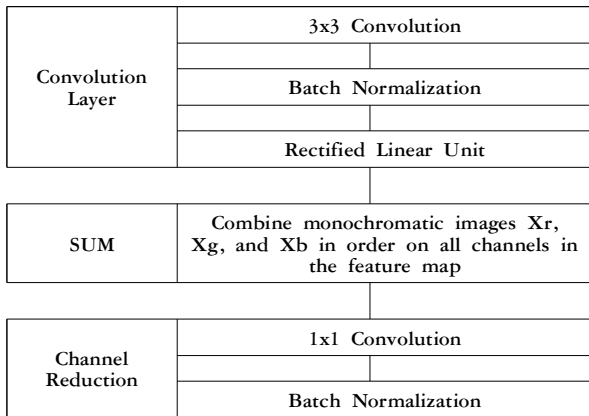


Fig. 2. Inner structure of RGB-csb

위와 같은 학습과정을 통한 RGB-csb는 모두 같은 영상 크기로 구성되며, 최종 결과물의 채널수는 3개이다.

Table 2. Test Accuracy

model	Test Accuracy (10 times)										avg.
	1	2	3	4	5	6	7	8	9	10	
CNN-160	61.88	62.50	61.87	60.00	58.75	60.00	59.38	61.25	58.52	61.25	60.54
rgbCNN-160	58.12	67.50	61.25	62.50	65.63	65.00	63.75	60.50	61.88	62.50	62.86
CNN-320	70.00	70.94	72.81	71.25	72.50	72.81	71.88	71.50	70.00	72.19	71.59
rgbCNN-320	74.06	74.06	73.13	74.37	70.00	72.81	72.25	75.62	71.87	73.13	73.13
CNN-640	79.53	78.91	81.09	81.41	83.28	82.97	82.66	81.56	80.78	81.41	81.36
rgbCNN-640	80.38	83.31	83.72	80.81	82.34	82.81	82.19	80.94	82.97	83.44	82.29
CNN-960	83.65	83.23	83.31	83.33	82.81	83.02	82.60	81.98	82.08	81.98	82.80
rgbCNN-960	84.27	84.15	83.02	84.73	84.06	83.32	83.40	83.08	82.46	83.88	83.64

### III. Training and Testing Images

제한된 영상에서의 성능 측정을 위하여 다음 Fig. 3과 같은 구조를 갖는 신경망을 사용하고, Table 1과 같은 방법으로 제한 영상에 대한 모델을 분류한다.

Table 1. Model classification table

Model	explanation
CNN-160	using 160 images by each category
rgbCNN-160	
CNN-320	using 320 images by each category
rgbCNN-320	
CNN-640	using 640 images by each category
rgbCNN-640	
CNN-960	using 960 images by each category
rgbCNN-960	

Table 1에서 제안된 모델명 뒤의 숫자는 카테고리 별 훈련 및 시험 영상 개수를 나타내며, 증가되는 개수는 그 전 모델에서 사용된 영상을 포함한다. CNN은 기존 CNN모델을 그대로 사용하는 것을 나타내며, Fig. 3.에의 RGB-csb 대신 컨볼루션층이 추가된 형태로 6개의 컨볼루션층과 2개의 완전연결층(fully connected layer)을 사용한다. rgbCNN은 본 논문에서 제안한 RGB-csb를 포함한 구조를 사용하는 것을 나타낸다.

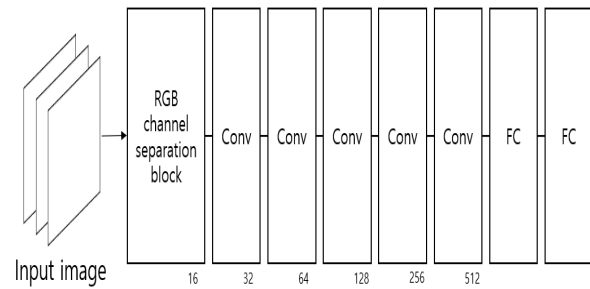


Fig. 3. CNN structure used in Experiments

## IV. Experiments

### 1. Training

실험을 위하여 사용한 데이터는 구글에서 검색한 동물 영상데이터이며, 실험 환경은 다음과 같다.

OS	windows 10 pro 64 bit
CPU	i7 - 7700
RAM	16G
GPU	NVIDIA TITAN Xp

실험에 사용한 모든 신경망의 하이퍼파라미터는 동일하며, 입력 영상은 224 x 224 픽셀이며 batch size는 32로 설정하고 40에폭 만큼 훈련시켰다. 또한 초기 학습 속도는 0.01로 설정했으며 10에폭 마다 5로 나뉘 학습 속도를 조절하였으며, RGB-csb를 포함하여 총 6개의 컨볼루션층과 2개의 fully connected layer를 사용하였다. 학습은 비교모델마다 각 10번씩 진행하였다.

### 2. Results

제안 모델의 영향을 세부적으로 파악하기 위하여 Table 2와 같이 다양한 조건의 모델들의 정확도 성능을 비교 분석한 결과, 몇 가지 의미 있는 결과들이 도출되었다. 학습 결과에 대한 정확도의 신뢰도를 확보하기 위하여 학습을 총 10번 수행하였다. 데이터가 많아질수록 비교적 평균 정확도가 높게 측정되었으나, 반면에 정확도의 증가폭이 데이터양이 적은 경우에 비하여 작아지므로 데이터양과 정확도의 상관관계가 로그함수로 유도됨을 알 수 있다.

rgbCNN-160의 경우 CNN-160보다 평균 정확도가 약 2.3% 높게 도출되었다. 이는 RGB-csb를 거치는 초기 입력 영상의 특징이 더욱 강조되어, 비록 편차에 의한 성능 저하 요인이 될 수 있으나 평균적으로 RGB-csb를 적용할 경우 유의미함을 알 수 있다.

rgbCNN-320, 640, 960의 경우 데이터가 많아질수록 각 학습 결과로 나타나는 정확도의 편차는 줄어들고, 또한 기존 CNN과의 정확도 성능 차이도 줄어든다.

## V. Conclusions

본 논문은 첫 번째 컨볼루션층을 RGB-csb로 대체하여, 입력 영상의 RGB 값을 특정 맵에 적용시켜 정확성을 제고시킬 수 있는 방법을 기반으로 훈련 및 시험 영상 수에 따른 정확도를 분석하였다.

rgbCNN모델을 기존의 CNN 모델과 학습 성능을 비교한 결과, 훈련 및 시험 영상 개수가 적을수록 rgbCNN모델이 약 2.3% 높은 정확도를 보였으나 영상 개수가 적을수록 각 학습 결과로 나타나는 정확도의 편차는 크게 나타났다. 반대로 영상의 개수가 많아질수록 기존 CNN과의 평균 정확도의 차이는 약 1%로 줄어들지만, 각 학습

결과와 정확도 편차는 줄어든다. 즉, 제안 방법에서 훈련 및 시험 영상 개수가 증가할수록 기존 CNN모델에 비하여 정확도의 차이가 적어지면서 로그함수의 상관관계를 나타냄을 알 수 있다.

결론적으로 실험 정확도가 전체적으로 향상된 것은 SUM 연산을 포함하는 RGB-csb를 입력 영상과 가장 가깝게 위치시켜 첫 번째 층에서 정확도 향상을 유도한 결과라고 판단된다.

## ACKNOWLEDGEMENT

이 논문은 2019년도 정부(산업통상자원부)의 재원으로 한국산업 기술진흥원의 산업전문인력역량강화사업(N0002428)과 정부(교육부)의 재원으로 한국연구재단의 기초연구사업(2018R1D1A1B07 051045)지원을 받아 수행된 것임

## REFERENCES

- [1] C. Szegedy, W. Lium, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke and A. Rabinovich, "Going deeper with convolutions", CVPR, pp.1894-1903, 2016
- [2] K. He, X. Zhang, S. Ren and J. Sun, "Deep Residual Learning for Image Recognition", CVPR, pp.770-778, 2016
- [3] G. Huang, Z. Liu, K. Q. Weinberger and L. Maaten, "Densely connected convolutional networks", CVPR, pp.4700-4708, 2017
- [4] A. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam. "Mobilenets: Efficient convolutional neural networks for mobile vision applications", arXiv preprint arXiv:1704.04861, 2017. 4
- [5] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting", The Journal of Machine Learning Research, vol.15(1), pp.1929-1958, June 2014
- [6] S. Ioffe and C. Szegedy. "Batch normalization: Accelerating deep network training by reducing internal covariate shift", In ICML, 2015
- [7] A. Krizhevsky, I. Sutskever and G. Hinton, "Imagenet classification with deep convolutional neural networks", Advances in neural information processing systems, pp.1097-1105, 2012

- [8] J. Kong, Y. Lee, M. Jang, "Convolution Neural Network with RGB channel separation block", Proc. of MITA2019, (In press)