

# 시각정보처리와 영상인식

조재현<sup>o</sup>

<sup>o</sup>부산가톨릭대학교 컴퓨터공학과

e-mail: jhcho@cup.ac.kr<sup>o</sup>

## Visual Information Processing and Image Recognition

Jae-Hyun Cho<sup>o</sup>

<sup>o</sup>Dept. of Computer Engineering, Catholic University of Pusan

### ● 요약 ●

기계학습은 문제가 복잡하여 수학적으로 정의를 하는 것이 어려울 때 유용하게 쓸 수 있는 방법으로 최근 패턴 또는 영상을 인식하기 위하여 급속도로 많이 사용되고 있다. 본 논문에서는 기존의 학습 모델과는 다르게 인간의 시각정보처리과정 중 망막의 특성과 시각피질의 특성을 고려한 모델을 제시하여 학습의 첫 단계인 원시 데이터를 가공하는 과정을 개선함으로써 좀 더 효율적인 인식모델을 제안하고자 한다.

**키워드:** 학습데이터(Learning data), 영상 압축(Image compression), 특징추출(Feature extraction)

### I. Introduction

기계 학습 과정 중 지도학습 방식은 크게 학습데이터 준비, 모델 준비, 모델 학습, 모델 평가 및 분석, 배포의 여러 단계로 구성되며 기계학습 과정은 개발하는 모델이 사람이 의도한 특정 정확도에 도달하기까지 각 단계별 조정을 가하면서 학습 과정을 지속적으로 반복해야하는 특징이 있다[1]. 무엇보다도 중요한 단계는 학습 데이터 구성 시 원영상의 특징을 이용하는 단계로서 좀 더 효율적으로 성능을 향상시킬 수 있다[1][2]. 본 논문에서는 영상 인식을 위하여 원 영상을 먼저 압축영상으로 변환한 후 인간의 시각시스템의 특성을 고려한 모델을 제안하고자 한다.

### II. Human Visual System and Image Recognition Model

인간의 시각체계는 광학계통, 망막 시각통로의 3단계로 나눌 수 있다. 1단계인 광학계통의 주된 기능은 입력되는 광선에너지에 담겨 있는 정보 즉 이미지가 망막위에 떨어 질 때 초점을 맞추어 주는 일이라 할 수 있다. 제일 먼저 각막에서 이미지를 형성하기 위하여 광선을 구부리는 기능을 가진 각막에서 이미지를 형성하여, 공막과 렌즈를 통해 2단계인 망막에 전달된다. 망막은 수용기와 뉴런을 모두 갖고 있어 시각정보의 변환이 종료됨과 동시에 처리가 시작되는 곳이다[3]. 망막세포에는 크게 수평세포(horizontal cell), 양극세포(bipolar cell), 아마크린 세포(amacrine cell), 신경절 세포(ganglion cell) 라 불리는 4개의 세포로 구성되어 있다. 그중에서 망막의 신경절

세포는 한 개의 눈에 대략 100만 개뿐이므로 1억 2천 5백만 개의 수용기로부터 신경절 세포에 이르기까지 상당한 양의 시각정보의 압축이 일어난다[3].

또한 신경절 세포는 망막의 신경 정보가 뇌로 전달되는 유일한 통로이므로 그 유형은 뇌로 입력되는 시각 정보의 유형을 결정하며[4], 망막영역의 마지막 출력에 해당하며 시각정보를 외측슬상핵(Lateral Geniculate neuron)으로 전달하는 역할을 한다. 외측슬상핵은 뇌의 왼쪽과 오른쪽에 하나씩 위치하고 있으며 약 150만개의 세포가 있으며, 매그노 세포 2개의 층과 파보 세포 4개의 층, 총 6개의 층으로 구성되어 있다. 각각의 외측슬상핵으로부터 대뇌 피질의 후두엽으로 시각정보가 들어가며, 이 후두엽의 시각 수용영역은 외측슬상핵보다 복잡하며 약 2억 5천만 개 이상의 신경세포들로 구성되어 있다. 이 수용영역을 선조피질이라고 한다[3]. 본 논문에서는 인간시각시스템에 기반하여 먼저 망막세포의 특성 중 압축 성질을 반영하고 시각피질의 특성인 방향성 반응에 대한 특성을 고려하여 인식 모델을 제안하고자 한다. 그림 1에 제안된 모델을 나타내었다. 학습 데이터를 구성하기 위하여 먼저 원영상을 압축영상으로 변화 후[5-7] 방향성에 민감한 특성을 고려하여 수직, 수평, 대각선, 역대각선 방향으로 각각 추출하여 학습 데이터를 구성하여 인식한다.

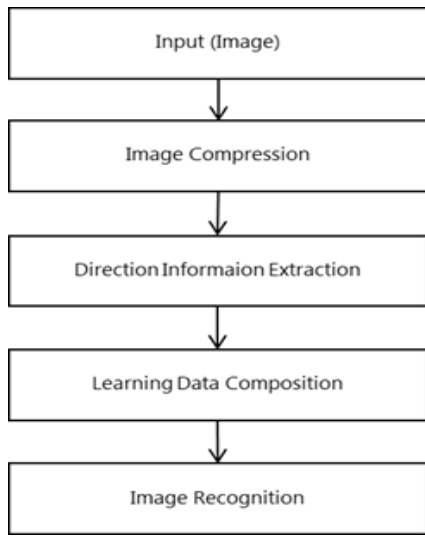


Fig. 1. Proposed model

### III. Experiment and Conclusions

실험에 적용된 영상은 64×64 크기를 가진 얼굴 영상 400개 중에서 280, 320, 360, 400개로 학습 데이터를 구성하였다[8]. 또한 다양한 응용 분야에서 널리 사용되는 오류 역전파 알고리즘을 개선한 Delta-bar-delta 알고리즘을 사용하였으며, 학습률은 0.01, 반복횟수는 10,000회이다. 학습한 결과는 표 1에 나타내었다. Range 블록 크기를 4 × 4, 그리고 8 × 8 로 하여 압축 정도를 변경하여 학습 데이터수에 따른 인식률을 비교한 결과를 나타내었으며 압축한 영상에 방향성의 정보를 사용하여 학습데이터가 많아짐에 따른 인식률이 높아짐을 알 수 있으며 차후 학습 데이터를 증가하고 또한 수직방향과 수평방향의 가중치의 변화를 줌으로써 좀 더 개선된 인식모델을 제시하여 성능을 개선하고자 한다.

Table 1. Recognition rate by learning data

Recognition rate by learning data(%)				
Size of Range block	The number of training data			
	280	320	360	400
4 × 4	88	90.5	95.5	99
8 × 8	88	91.5	95.25	99.25

### REFERENCES

[1] K. H. Lee, M K, Hwang, W. K. Sung, “Research Trends in Data Management Technology for Deep Learning”, Communications of the Korean Institute of Information Scientists and Engineers, vol. 37, no. 8, pp. 13-20, 2019.  
 [2] S. E. Moon, S. B. Jang, J. H. Lee, J. S. Lee, “Research Trends in Machine learning and Deep Learning”, The

Journal of The Korean Institute of Communication Sciences, vol. 33, no. 10, pp. 49-56, 2016.  
 [3] E. Bruce Goldstein, “Sensation & Perception”, Wadsworth, pp. 39~56, 2007.  
 [4] Baden T, Berens P, Franke K, Roson MR, Bethge M and Euler T, The function diversity of retinal ganglion cells in the mouse, Nautre 529:345-350, 2016  
 [5] Youngbong Kim, Yunjung Lee, “Fractal Compression using Range Block Coherence”, Journal of the Korean Institute of Information Scientists and Engineers, Vol. 27. No.2, pp. 117-122, Feb, 2000.  
 [6] Y. Fisher, E. W. Jacobs, and R. D. Boss, "Fractal Image Compression Using Iterated Transforms." Technical Report, Naval Ocean Systems Center, San Diego, CA92142 -5000.  
 [7] A. E. Jacquin, "Image Coding Based on a Fractal Theory of Iterated Contractive Image Transformations", IEEE Transaction on Image Processing , Vol 1, pp. 18-30, Jan, 1992.  
 [8] <https://www.cl.cam.ac.uk/research/dtg/attarchive/facedata.html>