영상인식을 위한 화질의 데이터 분류성

조재현O
O부산가톨릭대학교 컴퓨터공학과
e-mail: jhcho@cup.ac.krO

Data Classification of Visual Quality for Image Recognition

Jae-Hyun ChoO
ODept. of Computer Engineering, Catholic University of Pusan

● 요 약 ●

패턴 또는 영상을 인식하기 위하여 먼저 기계 학습 모델을 선택하고, 선택된 모델은 여러 단계의 처리 단계 과정으로써, 학습 데이터 구성과 특징 추출 그리고 분류기 등으로 크게 나눌 수 있다. 기존의 학습 모델의 처리 단계 중 학습 데이터 구성은 첫 번째 중요한 단계이다. 본 논문에서는 학습 데이터들의 특징을 분석하여 데이터 분류성의 척도로 사용될 수 있는지를 검토하여 차후 기계 학습 및 딥 러닝의 인식을 높이고자 한다.

키워드: 학습데이터(Learning data), 차원(Dimension), 특징추출(Feature extraction)

I. Introduction

지도학습 방식의 기계학습 과정은 크게 학습데이터 준비, 모델 준비, 모델 학습, 모델 평가 및 분석, 배포의 여러 단계로 구성된다[1]. 또한 기계학습 과정은 개발하는 모델이 사람이 의도한 특정 정확도에 도달하기까지 각 단계별 조정을 가하면서 학습 과정을 지속적으로 반복해야하는 특징이 있다 [1][2]. 본 논문에서는 영상인식을 위하여 압축영상을 분석하여 데이터 분류성의 척도로서 사용될 수 있는지를 제안하고자 한다.

II. Data classification of visual quality for image recognition

기계 학습(딥러닝) 처리 과정은 일반적으로 그림 1과 같이 입력 데이터를 학습데이터를 구성하고 특징추출과정을 거쳐 학습알고리즘을 수행한 후 테스트 데이터를 이용하여 일치하는지를 확인하는 테스트 과정으로 구분할 수 있다.

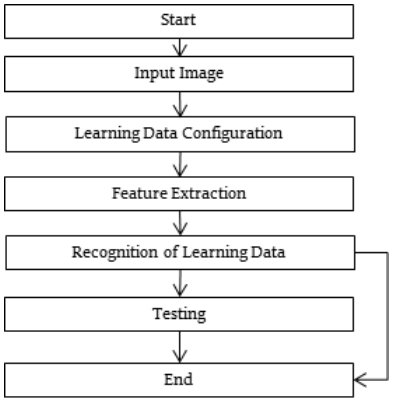

Fig. 1. Steps of Supervised Machine Learning

기계 학습 과정에서 가장 중요한 것 중 하나는 주어진 문제를 풀기에 적합한 특징을 추출하는 것이다. 주어진 학습데이터를 그대로 사용하는 것도 가능하지만 필요한 정보 이외에도 관련 없는 다른 정보나 노이지 등이 포함되어 있기 때문에 데이터를 가공 없이 사용하는 것은 오히려 모델의 성능을 저하 시킬 수 있다[2]. 일반적으로 주어진 데이터는 높은 차원을 가지는 경우가 많기 때문에 보다 낮은 차원의 특징으로 변환할 필요성이 있다. 데이터의 차원이 증가할수록 상대적으로 주어진 데이터의 희소성도 증가하게 된다. 즉, 출력 값과 데이터 간의 관계를 학습하기 위한 데이터가 부족한 현상이 나타난다.

또한 학습 대상이 되는 파라미터의 개수도 증가하기 때문에 총체적으로 학습에 더 많은 양의 데이터가 요구되는 결과를 야기한다[3]. 본 논문에서는 영상 인식에 중요한 데이터의 특성 중 데이터의 분류성의 척도를 검정하기 위하여 그림 2와 같이 영상의 자기유사성의 특징을 가지고 차원의 따른 영상으로 변환하여 학습 데이터로 재구성하였다. 여러 압축 기법 중 프랙탈 압축은 영상의 일부분은 같은 영상의 다른 영역과 모양이 거의 유사하다는 자기 유사성에 기초한 방법으로 두 블록간의 수렴하는 아핀 변환을 찾는 작업으로 다시 디코딩 작업을 통해 압축된 영상을 구할 수 있다[4].

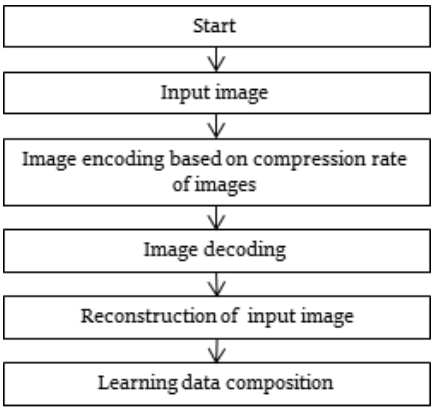

Fig. 2. Data composition for image recognintion

III. Experiment and Conclusions

실험에 적용된 영상[5]은 64×64 크기를 가진 얼굴 영상 400(40명 X 10)개에 대한 영상 화질을 비교하였다. 영상에 대한 화질의 데이터 분류성이 가능한지를 분석하기 위하여 T-검정을 통하여 분석하였다. 같은 차원 즉 압축률이 같은 조건에서 유사 영상 간에 화질의 정도가 차이가 있음을 나타내고자 한다. 여러 영상 클래스 중 에서 2개의 영상 클래스에 대한 영상 화질은 다음과 같다. Table 1은 여러 영상 중 2개의 클래스들(각 10개의 영상)에 대한 영상의 화질을 PSNR(Peak Signal-to-Noise Ratio)로 나타내었으며 Table 2는 10개의 영상에 대한 각 클래스의 T검정 결과와 신뢰구간을 나타내었다. 입력 영상선택 시 신뢰구간을 이용할 수 있으며 같은 척도 즉 같은 압축률에서 영상에 대한 화질을 영상인식을 위한 입력 데이터들로 구성할 수 있음을 알 수 있었다.

Table 1. PSNR of reconstructed input image(sample)

PSNR		
image	class 1	class 2
1	26.28380	24.45945
2	26.33455	24.42457
3	25.91852	23.84423
4	27.30772	24.81753
5	25.72841	23.99693
6	27.33579	24.43729
7	26.94654	23.78630
8	27.20958	23.39605
9	26.64796	24.10605
10	26.10748	23.95648

Table 2. Confidence interval of 2 classes

class	Significance level (0.05)		
	p_value	lower	upper
image1	0.0000	26.1575	27.0066
image2	0.0000	23.8257	24.4192

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] Hughes, G. “On the mean accuracy of statistical pattern recognizers,” IEEE Transactions on Information Theory, vol. 14, no. 1, pp. 55-63, 1968.

[4] Y. B. Kim, Y. J. Lee, “Fractal Compression using Range Block Coherence”, Journal of KISE: Computer Systems and Theory, vol. 27, no. 2, pp. 117-122, 2000.

[5] *https://www.cl.cam.ac.uk/research/dtg/attarchive/facedata.html*