

몰포로지 신경망 기반 딥러닝 시스템

최 종 호*

Deep Learning System based on Morphological Neural Network

Jong-Ho Choi*

요 약 본 논문에서는 몰포로지 연산을 기본으로 하는 몰포로지 신경망(MNN: Morphological Neural Network) 기반 딥러닝 시스템을 제안하였다. 딥러닝에 사용되는 레이어는 몰포로지 레이어, 풀링 레이어, ReLU 레이어, Fully connected 레이어 등이다. 몰포로지 레이어에서 사용되는 연산은 에로전, 다이레이션, 에지검출 등이다. 본 논문에서 새롭게 제안한 MNN은 기존의 CNN(Convolutional Neural Network)을 이용한 딥러닝 시스템과는 달리 히든 레이어의 수와 각 레이어에 적용되는 커널 수가 제한적이다. 레이어 단위 처리시간이 감소하고, VLSI 칩 설계가 용이하다는 장점이 있으므로 모바일 임베디드 시스템에 딥러닝을 다양하게 적용할 수 있다. MNN에서는 제한된 수의 커널로 에지와 형상검출 등의 연산을 수행하기 때문이다. 데이터베이스 영상을 대상으로 행한 실험을 통해 MNN의 성능 및 딥러닝 시스템으로의 활용 가능성을 확인하였다.

Abstract In this paper, we propose a deep learning system based on morphological neural network(MNN). The deep learning layers are morphological operation layer, pooling layer, ReLU layer, and the fully connected layer. The operations used in morphological layer are erosion, dilation, and edge detection, etc. Unlike CNN, the number of hidden layers and kernels applied to each layer is limited in MNN. Because of the reduction of processing time and utility of VLSI chip design, it is possible to apply MNN to various mobile embedded systems. MNN performs the edge and shape detection operations with a limited number of kernels. Through experiments using database images, it is confirmed that MNN can be used as a deep learning system and its performance.

Key Words : CNN, Deep Learning, Embedded System, MNN, Morphology, VLSI

1. 서 론

현재 딥러닝 시스템으로 널리 활용되고 있는 CNN(convolutional neural network) 분야[1,2,3]에서 일반적으로 고려되고 있는 파라미터는 레이어 수, 각 레이어에서의 노드 수, 레이어간 노드의 연결 경로, 신호전달 체계에서의 활성화 함수 등이다. 이러한 파라미터들의 설정에는 복잡한 연산을 필요로 함에도 불구하고, 효과적인 자동설정 알고리즘이 개발되어 있지 않았으므로 수동으로 설정하는 경우가 대부분이었다. 파라미터 설정의 문제는 뉴럴 네트워

크에서의 레이어 수 즉 히든 레이어의 수를 증가시키는 것을 불가능하게 만든 큰 원인이 되었고, 이로 인해 레이어를 깊게 설계하여도 뉴럴 네트워크의 성능개선은 크게 일어나지 않는다는 암흑의 세계로 빠져 들어갈 수밖에 없었다. 이러한 예측은 당시의 컴퓨팅 파워에도 큰 관련이 있었다.

현재는 인터넷으로 대표되는 네트워크의 발전으로 대량의 입력데이터가 자연스럽게 확보되고, CPU와 함께 GPU로 대표되는 컴퓨팅 파워의 급격한 발전에 힘입어 히든 레이어의 수를 증가시킨 딥러닝 방식의 머신러닝이 활발하게 응용되고 있다. 비지도 학습의

This Research was Supported by Kangnam University Research Grants(2017).

*Department of IoT Electronic Engineering, Kangnam University(jhchoi@kangnam.ac.kr)

Received January 18, 2019

Revised February 16, 2019

Accepted February 20, 2019

형태로 발전하고 있는 딥러닝에서는 224 × 224 입력영상에 적용되는 레이어의 수가 100여개 이상으로 증가하고 있고, 레이어 단위 커널 수도 100여개 이상이 사용되고 있다. 기술적 진보의 차원에서 빅 데이터와 GPU 기반 컴퓨팅 파워를 근간으로 딥러닝을 설계하는 방향으로 머신러닝이 진화하고 있으나, 학습이 완료된 경우라도 컴퓨팅 파워의 부담은 커질 수밖에 없어 그 응용이 제한되고 있다.

한편 기술적 응용의 차원에서 스마트 폰으로 대표되는 임베디드형의 모바일 시스템 또는 지능형 CCTV 등 기타 소형 시스템에 딥러닝 알고리즘을 하드웨어 또는 소프트웨어의 형태로 탑재하고자 하는 새로운 산업적 수요가 발생하고 있다. IoT, 네비게이션 시스템, CCTV 기반 ITS 및 방범방재 시스템, 스마트 폰 앱 등이 대표적인 예이다. 그러나 이러한 분야에서 컨볼루션 레이어, 풀링(pooling) 레이어, FC(fully connected) 레이어 등 다수의 레이어로 구성되는 현재의 딥러닝 시스템을 구성하는 것은 불가능하다.

따라서 본 논문에서는 물포러지 연산(morphological operation)을 기본으로 하는 물포러지 신경망 즉 MNN(morphological neural network) 기반 딥러닝 시스템을 제안하였다. MNN에서의 기본 연산은 에로전(erosion), 다이레이션(dilation), 오픈닝(opening), 크로우징(closing)으로 연산과정이 단순하고 디지털 칩 설계가 용이하다. 현재의 CNN(convolutional neural network) 기반 딥러닝 시스템과는 달리 MNN 딥러닝 시스템의 장점은 처리시간 감소 및 VLSI 칩 설계를 통해 컴퓨팅 파워에 의존할 수밖에 없는 모바일 임베디드 시스템에 딥러닝 기법을 다양하게 활용할 수 있다는 것에 있다.

2. MNN 딥러닝 시스템

현재 기계학습의 일환으로 널리 각광받고 있는 CNN 기반 딥러닝 시스템에서의 특징맵 추출 과정을 그림 1에 나타냈다.

CNN에서 시스템의 성능을 좌우할 정도의 중요한 연산은 컨볼루션이다. 데이터를 입력하면 다단계의

레이어에서 특징을 추출하고 결합하는 역할을 하는 것이 컨볼루션이다. 입력 레이어에 근접한 부분에서는 low level의 특징에 해당하는 에지, 텍스처 등을 검출하고, 출력 레이어에 근접할수록 high level의 특징에 해당하는 형상이 추출된다. 특징추출 과정에서 컨볼루션은 저역통과필터, 에지검출필터, 고역통과필터 등의 역할을 수행하는 것이다.

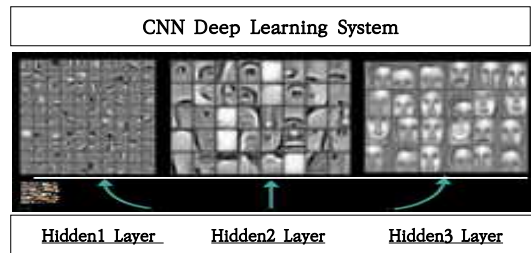


그림 1. CNN 구조
Fig. 1. CNN Architecture

컨볼루션에 해당하는 물포러지 기본연산은 다이레이션, 에로전 다이레이션-에로전, 에로전-다이레이션이다. 본 논문에서 새롭게 제안한 MNN도 CNN과 동일한 개념의 접근방식이다. 따라서 물포러지 연산을 사용하여 에지검출, 형상검출, 저역통과필터, 고역통과필터의 역할을 수행하여야 한다. 특징추출에 해당하는 물포러지 연산식은 다음과 같다[4,5,6].

$$\text{Erosion} : I \ominus B = \bigcap_{w \in B} I_w \quad (1)$$

$$\text{Dilation} : I \oplus B = \bigcup_{w \in B} I_w \quad (2)$$

$$\begin{aligned} \text{External Edge} & (I \oplus B) - I \\ \text{Internal Edge} & I - (I \ominus B) \end{aligned} \quad (3)$$

$$\text{Morphological Edge} (I \oplus B) - (I \ominus B)$$

$$\text{Hit/Miss} : f \otimes B = (f \ominus B_1) \cap (f^c \ominus B_2) \quad (4)$$

$$S_n(X) = [(X \ominus nB) \setminus [(X \ominus nB) \circ B]] \quad (5)$$

$$\text{Peak}(f) = f - (f \circ B) \quad \text{Valley}(f) = (f \bullet B) - f \quad (6)$$

여기서 \setminus 는 차집합 연산자이고, nB 는 B 를 n 회 다이

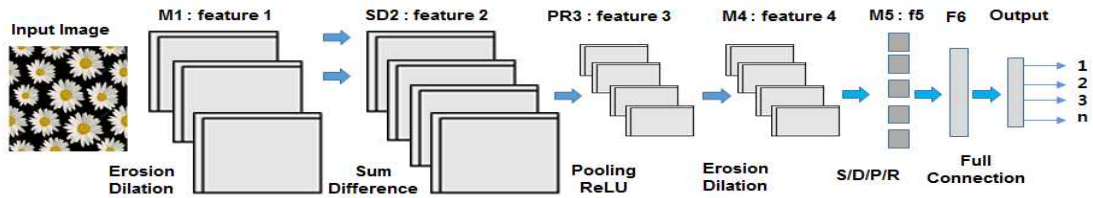


그림 2. 제안 MNN 구조
Fig. 2. Proposed MNN architecture

레이션 연산한 결과이다.

그림 2에 본 논문에서 제안한 MNN의 구조를 나타냈다. MNN에서는 CNN에서의 컨볼루션 연산과 유사하게 다일레이션과 에로전 연산을 기본으로 에지, 스킴리톤 등을 추출하고, 이를 풀링(Pooling), ReLU(Rectified Linear Unit) 등과 연결하여 차원을 축소하며, 최종적으로는 FC(Fully connected) 레이어에서 인식 결과를 출력하는 구조이다.

그림 2에 나타난 MNN 구조는 기본적인 형태로 몰포리지 연산을 반복 적용함으로써 레이어의 수를 증가시키는 형태로 변형할 수 있다. 몰포리지 연산을 이용한 MNN 시스템은 CNN과 동일하게 다양한 형태로 구성할 수 있다. MNN은 새롭게 제안하는 딥러닝 시스템 방식이라는 측면에서 기본적인 구성을 나타냈다. 레이어의 수와 각 레이어에 적용되는 커널의 종류, 몰포리지 연산 레이어 등은 시스템 설계자가 시스템의 특성을 고려하여 재배열 할 수 있다.

컴퓨터 비전 분야에서 널리 이용되고 있는 딥러닝에서의 입력은 컬러영상으로 R/G/B 영상으로 분해하면 각각은 그레이스케일 영상이되고, 각각은 딥러닝 시스템의 입력이 된다. 컬러값을 특징으로 사용하지 않는 분야에서는 R/G/B 화소값의 평균을 취하는 방식으로 얻은 그레이스케일 영상을 입력으로 할 수도 있다.

입력영상으로부터 특징을 추출하기 위한 몰포리지 연산에서는 일반적으로 3×3 또는 5×5 커널이 사용된다. 각각의 커널 크기에서 커널의 형태는 서로 다른 유일한 특징을 추출하는 것을 목표로 다양하게 구성할 수 있다. 3 × 3인 경우에는 512개의 커널이, 그리고 5 × 5인 경우에는 2²⁵개 정도의 많은 커널이 대상이 된다.

CNN의 컨볼루션 레이어에 해당하는 첫번째 레이어에서는 에로전과 다일레이션 연산을 수행한다. 최근 들어 딥러닝에서의 특징추출 과정을 시각화하는 연구가 진행되면서 입력에 가까운 레이어에서는 수평, 수직 에지와 같은 로컬한 특징이 추출되고, 컨볼루션 및 풀링 과정을 거치면서 출력에 가까운 레이어에서는 글로벌한 형상 특징이 추출되며, 최종적으로 FC 레이어에서는 입력영상을 대표하는 강력한 특징이 추출되는 것으로 해석되고 있다.

에로전과 다일레이션의 합 또는 차 연산을 수행하는 두번째 레이어에서는 영상의 강력한 로컬 특징인 에지와 스킴리톤을 구하는 연산을 수행한다. 이 부분은 CNN에서 컨볼루션 커널의 웨이트 설정에 관한 부분이다. 또한 특징맵의 차원을 감소시키고 토폴로지 변화에 의존하지 않는 특징을 얻기 위해 다일레이션-에로전-sum/difference-풀링-ReLU 연산을 수행하는 레이어를 히든 레이어로 설계한다. 다수의 히든 레이어를 거쳐 특징맵의 차원이 축소되면서 최종적으로는 FC 레이어에서 개체를 인지하는 강력한 특징을 얻을 수 있다. 이와 같이 MNN은 몰포리지 연산 레이어, 풀링(서브샘플링) 레이어, FC 레이어로 구성되는 시스템으로 몰포리지와 풀링 연산을 다수의 레이어로 설계하는 딥러닝 기법이다. 여기서 풀링, ReLU, FC 레이어는 CNN과 동일한 연산이다.

MNN에서의 또 다른 주제는 학습과정이다. CNN에서의 학습은 커널의 계수를 최적으로 설정하는 것이다. 이에 반해 MNN에서의 학습은 최적의 커널을 선택하는 과정이다. CNN에서의 커널이 가질 수 있는 웨이트 값의 개수는 무한대이다. MNN에서는 이에 대응되는 방식으로 각 계층에서 사용되는 커널의 종류와 바이어스 값을 결정하는 것이다. 응용분야에 따

라 하나의 상수 값을 바이어스로 할 수도 있고, 3×3 또는 5×5 커널에서 각 요소 단위로 바이어스를 결정하는 방법으로 확장할 수도 있다. 본 논문에서 제안한 MNN에서는 바이어스 값은 0으로 고정하고, 커널의 종류만을 선택하는 방식이다. 사용할 수 있는 커널의 개수는 유한하다.

ImageNet LSVRC에서 우승하고 현재에도 CNN의 기본 모델로 활용되고 있는 AlexNet은 5개의 컨볼루션 레이어로 구성된다. 5개의 컨볼루션 레이어 중 첫번째 컨볼루션 레이어에서는 11×11 크기의 96개 커널, 이후 컨볼루션 레이어에서는 각각 5 × 5 256커널, 3×3 384커널, 3×3 384커널, 3×3 256 커널이 사용된다. 대규모 영상인식 모델이기 때문에 높은 컴퓨팅 파워를 필요로 한다. 그러나 본 논문에서 새롭게 제안한 MNN은 임베디드 시스템에 적용되는 소규모 물체인식 용도의 모델이기 때문에 CNN 컨볼루션 레이어에 해당하는 각각의 물포러지 레이어에서 사용되는 커널 수는 그림 3에 나타난 square, circle 등 최대 20여개만이 사용된다.

응용분야에 따라서 상이하나 커널의 종류는 크게 줄일 수 있다. 3×3 커널인 경우, 활용가능한 커널의 개수는 2⁹=512개이다. 그러나 커널영역이 3개 미만인 경우를 제외하면 커널의 수는 크게 감소한다. 레이어의 수 또한 비교적 적기 때문에 두개의 계층을 서로 묶어 역전파 방식의 학습을 무작위로 수행하는 것이 가능하다. 그림 3에 MNN에서 사용되는 커널을 나타냈다.

입력영상으로부터 특징을 추출하기 위한 물포러지 연산에서는 일반적으로 3×3 또는 5×5 커널이 사용된다. 각각의 커널 크기에서 커널의 형태는 서로 다른 유일한 특징을 추출하는 것을 목표로 다양하게 구성할 수 있다. 그림 3에 사용될 수 있는 커널의 종류를 나타냈다.

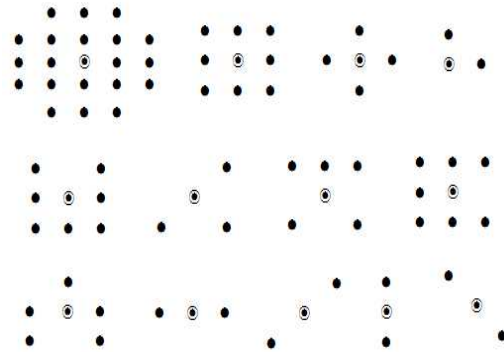


그림 3. MNN 커널
Fig. 3. MNN Kernel

3. 실험

데이터베이스 형태로 공개된 꽃영상을 대상으로 본 논문에서 제안한 MNN 딥러닝 시스템의 유용성을 검증하였다. 그림 4에 본 실험에서 사용한 실험영상 일부를 나타냈다.



그림 4. 테스트 영상(daisy,dandelion,rose,tulips)
Fig. 4. Test image(daisy,dandelion,rose,tulips)

MNN 시스템의 첫번째 레이어인 물포러지 레이어에서의 연산 결과를 그림 5에 나타냈다. 매트랩에서 3×3 Square 필터를 사용하여 수행한 다이레이션과 에로전 연산 결과이다[7].



(a) Dilation image (b) Erosion image

그림 5. 에로전과 다이레이션 영상

Fig. 5. Erosion and Dilation image

CNN 딥러닝 시스템에서 몰포리지 레이어 다음의 레이어는 sum/difference 레이어이다. 에지, 스케리톤, peak/value 등의 특징을 얻을 수 있는 레이어이다. 이러한 특징을 얻는 과정에서 사용되는 커널은 1 개 또는 다수로 고정시킬 수도 있고, 학습과정을 거쳐 선택할 수도 있다. 매트랩에서 커널의 변경에 따른 특징의 변화를 분석한 결과를 그림 6에 나타냈다. 커널의 선택이 다양한 특징의 변화를 유발하기 때문에 CNN에서와 동일하게 학습을 통해서 커널을 선정하는 것이 필요하다.

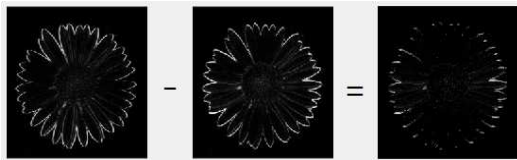
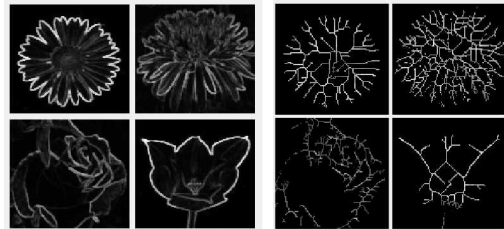


그림 6. 커널에 따른 특징 차이

Fig. 6. Feature difference by kernel

딥러닝 시스템의 경우, 최종 출력으로 향하는 과정에서 필수적인 것은 차원축소이다. 차원축소의 방법으로 Max-Pooling을 사용하였다. 1차 Max-Pooling은 2x2 필터를 적용하였으므로 224x224 영상이 112x112로 축소된다. 선형함수인 ReLU를 activation 함수로 사용한 Pooling-ReLU 연산 결과를 그림 7에 나타냈다. 이상과 같은 3개의 레이어에서 수행한다이레이션-에로전-Sum/Difference-Pooling/ReLU 연산은 딥러닝 시스템 설계 단계에서 동일한 형태로 반복하거나 변경하여 반복할 수 있다[7].



(a) Edge feature (b) Skeleton feature

그림 7. Pooling-ReLU 영상

Fig. 7. Pooling-ReLU image

다음 과정은 FC(Fully connected) 레이어이다. 최종출력의 노드로 차원을 축소하는 과정이다. FC 레이어의 구조를 그림 8에 나타냈다.

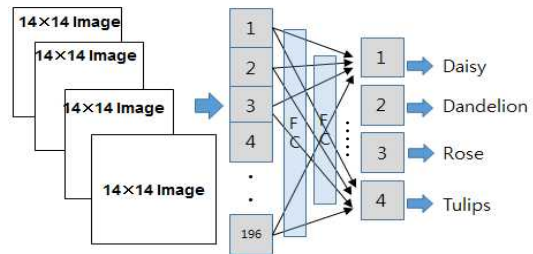


그림 8. Fully connected 레이어

Fig. 8. Fully connected layer

본 논문에서 제안한 MNN에서의 FC 레이어는 CNN에서 사용되는 것과 동일하게 적용되는 경우이므로 웨이트 학습방식에 관한 부분은 연구대상으로 하지 않았다. FC 레이어는 단일 또는 다수의 레이어로 설계할 수 있다. FC 레이어에서 사용되는 1x1 컨볼루션, dropout, softmax 등은 CNN에서 적용하는 것과 동일하게 설정할 수 있다.

그림 9에 나타낸 구글의 인셉션(inception) 뉴럴 네트워크[8]와 MNN 영상인식 결과를 비교하는 과정을 통해 본 논문에서 제안한 MNN의 성능과 딥러닝 적용 가능성을 평가하였다.

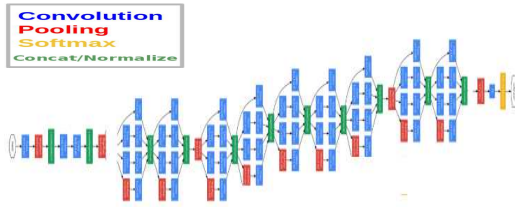




그림 9. 구글 인셉션 구조
Fig. 9. Google Inception Architecture

구글 인셉션 모델에서는 텐서플로우로 개발된 영상 인식 모듈[9]을 사용하였고, MNN에서는 매트랩으로 개발한 인식 모듈을 사용하였다. daisy, dandelion, rose, tulips, sunflower로 구성되는 꽃 데이터베이스에서의 구글 인셉션 모듈과 MNN의 인식결과 일부를 표 1에 나타냈다.

표 1. CNN(구글 인셉션 모듈)과 MNN의 인식 결과
Table 1. Recognition results in CNN(Google inception module) and MNN

| Image | Recognition Results | | |
|---|---------------------|--------|-------|
| | flower | CNN | MNN |
|  | daisy | 0.9835 | 0.970 |
| | dandelion | 0.0042 | 0.005 |
| | rose | 0.0003 | 0.001 |
| | tulips | 0.0013 | 0.004 |
| | sunflower | 0.0108 | 0.020 |
|  | daisy | 0.0204 | 0.120 |
| | dandelion | 0.0135 | 0.040 |
| | rose | 0.6462 | 0.590 |
| | tulips | 0.3132 | 0.230 |
| | sunflower | 0.0067 | 0.020 |

인식결과 커널수와 레이어가 다수인 CNN이 인식 대상이 되는 꽃의 가중치에서 MNN보다 비교적 높았으나, 그 결과는 인식에 큰 영향을 미치지 않는다. daisy, dandelion, rose, tulips, sunflower 각에서의 인식률은 구글 넷과 MNN 모두에서 97% 이상이었다. 확률이 가장 높은 출력을 선택하는 one-hot 변환으로 인식하였다. 본 실험은 CNN과 동일하게 임베디드 시스템에서 MNN의 딥러닝 적용

가능성을 판단하는 것이 목표이므로 비교적 단순한 5종류의 꽃을 분류하는 것이었다. 본 논문에서 새롭게 제안한 MNN 시스템이 임베디드 딥러닝 시스템으로 널리 활용될 수 있다는 것을 확인한 실험결과이다.

4. 결 론

본 논문에서는 MNN 기반 딥러닝 시스템을 새롭게 제안하였다. CNN의 중간 레이어에서 수행하는 컨볼루션을 물포러지 연산으로 대체하고, 다이레이션-에로전, Sum/ Difference, Pooling/ReLU 레이어에서의 연산을 통해 에지에서 형상으로 변하는 물체의 특징을 추출하고 인식하는 딥러닝 방식이다.

본 논문에서 제안한 MNN은 CNN과 상이하게 레이어의 수는 물론 레이어 단위 커널의 수가 제한적이기 때문에 VLSI 칩 설계가 용이하다는 특징을 갖는다. 따라서 모바일 임베디드 시스템 등에 딥러닝을 다양하게 활용할 수 있다는 장점이 있다.

꽃영상을 대상으로 수행한 실험을 통해 MNN 시스템의 유용성을 검증한 결과, 임베디드 딥러닝 시스템으로의 활용 가능성을 확인하였다. 향후의 연구과제는 커널 종류의 선택 및 MNN FC 레이어에서의 가중치 학습에 관한 최적의 알고리즘을 개발하는 부분이다.

REFERENCES

- [1] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks", *Advances in Neural Information processing Systems 25*, NIPS, 2012.
- [2] H. Lee, R. Grosse, R. Ranganath, and A.Y. Ng., "Convolutional deep belief networks for scalable unsupervised learning of hierarchical representations", *Proceedings of the 26th Annual International Conference on Machine Learning*, ACM, 2009.
- [3] Y. LeCun, K. Kavukcuoglu, and C. Farabet, "Convolutional networks and applications in vision", *International Symposium in Circuits*

and Systems (ISCAS), IEEE, 2010.

[4] Serra, J., *Image Analysis and Mathematical Morphology*, Vol.1, Academic Press, New York, 1982.

[5] Serra, J., "Introduction to Mathematical Morphology," *Computer Vision, Graphics, and Image Processing*, Vol.35, No.3, 1986.

[6] Jongho, Choi and Jongsoo, Choi, "A Morphological Recognition Algorithm of Blob Shaped Particles for Automated Visual Inspection", *Journal of KIISE*, Vol. 21, No.5, May, 1994.

[7] Jong-Ho Choi, "Morphological Object Recognition Algorithm", *Journal of KIIECT*, Vol. 11, No. 2, April, 2018.

[8] Szegedy, Christian, et al. "Going deeper with convolutions." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2015.

저자약력

최 종 호(Jong-Ho Choi)

[중신회원]



- 1982년 2월 : 중앙대학교 전자공학과(공학사)
- 1984년 2월 : 중앙대학교 대학원 전자공학과(공학석사)
- 1987년 2월 : 중앙대학교 대학원 전자공학과(공학박사)
- 1990년 3월 ~ 현재 : 강남대학교 IoT전자공학과 교수

〈관심분야〉

영상처리, 컴퓨터시각, 딥러닝