

컬러 영상 데이터베이스를 활용한 적외선 영상 객체 분류 신경망 학습 방법에 관한 연구

*강주성, *한현준, *Rohit Thakur, **김낙우, **이병탁, *이흥노

*광주과학기술원, **한국전자통신연구원

*k92492@gist.ac.kr, *hyunjun0701@gist.ac.kr, *trohit920@gist.ac.kr, **nwkim@etri.re.kr, **bytelee@etri.re.kr, *heungno@gist.ac.kr

Augmentation of RGB Training dataset for Improved Performance on Infrared Object Classifier

*Jusung Kang, *Hyunjun Han, *Rohit Thakur, **Nakwoo Kim,

**Byongtak Lee, and *Heung-No Lee

*Gwanju Institute of Science and Technology

**Electronics and Telecommunications Research Institute

요약

본 논문에서는 적외선 영상 객체 분류 신경망 학습을 위해 컬러 영상을 활용하는 방법에 대해 기술한다. 다섯 개의 선택된 목표 객체에 대해 ImageNet에서 제공하는 컬러 영상들을 수집하고, 인위적 패턴을 더하여 적외선 영상과 유사하게 만들고자 하였다. 이러한 컬러 영상들과 인위적 패턴이 더해진 영상들을 이용하여 GoogLeNet을 fine-tuning 하였으며, 적외선 영상을 분류함으로써 성능을 확인하였다. 실험 결과, 컬러 영상을 학습시켰을 때 보다 8.6%의 성능 향상을 확인하였으며, 이를 통해 인위적 패턴이 더해진 컬러 영상을 활용한 적외선 영상 객체 분류 신경망 학습 방법에 대한 가능성을 확인하였다.

I. 서론

지금까지 영상 신호처리 분야에서 객체 분류 기술은 주로 컬러 영상에 대해 연구되어 왔다. 컬러 영상의 경우 데이터 획득에 대한 용이성으로 인해 Pascal VOC[1]나 ImageNet[2]과 같은 대규모 영상 데이터베이스 구축 프로젝트가 진행되어 왔으며, 성능 평가 대회가 개최되어 다양한 기술들이 경쟁하듯 개발되어 왔다. 최근에는 GoogLeNet[3], VGGNet[4]과 같은 합성곱 신경망(Convolutional Neural Network)을 이용한 객체 분류 시스템이 개발되어 96.9%의 분류 정확도를 보여주었다[3].

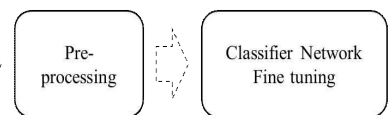
적외선 영상의 경우 기술의 발전으로 인해 최근에서야 데이터 획득이 용이해졌으며, 다양한 연구 그룹에서 OTCBVS Benchmark Dataset[5]과 같은 데이터베이스들을 공개하고 있다. 하지만 컬러 영상 데이터베이스와 비교해 보았을 때, 객체 수나 객체별 영상 수, 동일 객체 내의 객체 다양성 등 데이터베이스 규모 측면에서 현저히 부족한 것이 사실이며, 이를 활용하여 컬러 영상의 경우와 같은 일반적인 적외선 영상 객체 분류 신경망을 연구하기에는 어려움이 존재한다.

본 논문에서는 위와 같은 문제점을 해결하기 위해 컬러 영상을 활용한 효율적인 적외선 영상 객체 분류 신경망 학습 방법에 대해 기술한다. 목표 객체에 대한 컬러 영상들을 수집하고 인위적 패턴을 더하여 적외선 영상과 유사한 영상들을 만들었으며, 해당 영상들을 학습시켰을 때의 적외선 영상 객체 분류 성능을 비교함으로써, 제안하는 방법이 단순한 컬러 영상을 학습시켰을 때 보다 적외선 영상 객체 분류 신경망 학습에 효율적임을 확인하였다.

II. 본론

<Training phase>

RGB Object dataset



<Testing phase>

IR Object test dataset

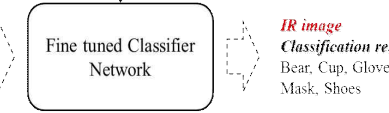


그림 1. 컬러 영상 기반 적외선 영상 객체 분류 신경망

본 논문에서 고려하는 컬러 영상을 활용한 적외선 영상 객체 분류 신경망 학습 방법은 그림 1과 같다. 먼저 공개되어있는 컬러 영상 데이터베이스로부터 목표 객체에 대한 컬러 영상들을 수집하고, 전처리 과정을 통해 인위적 패턴을 더하여 유사 적외선 영상들을 구성한다. 이렇게 구성된 영상들을 활용하여 신경망 fine-tuning 과정을 수행함으로써 학습을 완료한다. 다음으로 tuning된 신경망을 활용하여 입력된 테스트 적외선 영상을 분류하고 그 성능을 확인함으로써 제안하는 방법에 대한 학습 효율을 확인한다.

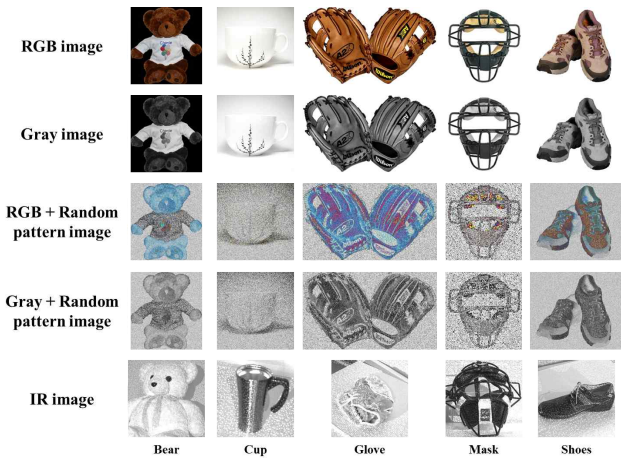


그림 2. 컬러·적외선 영상 데이터베이스

신경망 학습을 위해 준비된 영상 데이터베이스는 그림 2와 같다. 먼저 다섯 개의 목표 객체를 주변에서 쉽게 구할 수 있는 곰 인형, 컵, 야구 글러브, 야구 마스크, 신발로 선택 하였다. 다음으로 컬러 영상 데이터베이스의 경우 선택된 목표 객체들에 대해 ImageNet으로부터 총 6376장의 영상들을 수집하여 구축하였으며, 적외선 영상 데이터베이스의 경우 적외선 광원 및 수신기를 사용하여 객체별 100장씩 총 500장의 영상을 촬영해 구축하였다.

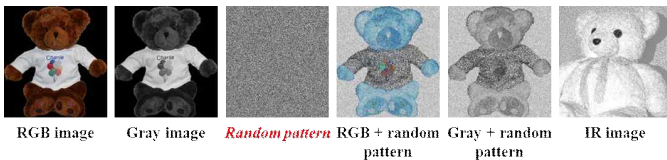


그림 3. 전처리 과정을 통한 유사 적외선 영상 데이터베이스

전처리 과정을 통한 유사 적외선 영상 데이터베이스를 구축하는 과정은 다음과 같다. 먼저 Gray scale을 지니는 적외선 영상의 특성을 고려하기 위해 컬러 영상을 회색 영상으로 변환하여 회색 영상 데이터베이스를 구축하였다. 다음으로 적외선 영상 속 광원으로 인해 나타나는 Speckle pattern을 고려하기 위해 [0, 255] 범위의 랜덤 패턴을 영상에 더함으로써 컬러 + 랜덤 패턴, 회색 + 랜덤 패턴 데이터베이스를 구축하였다.

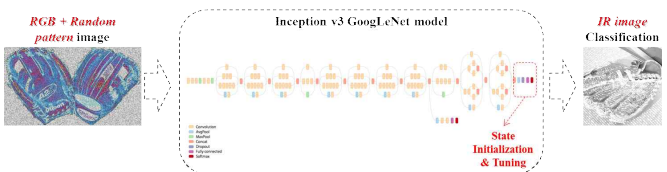


그림 4. GoogLeNet 기반 Fine tuned Network

본 논문에서 고려되는 객체 분류 신경망은 ImageNet Challenge 2014 대회 객체 분류 분야에서 우승한 GoogLeNet이다. GoogLeNet은 Local Reception Field 안에서의 Non-linearity 특성을 좀 더 잘 추출하기 위해 다양한 크기의 합성곱 필터를 사용한 Inception v1 module을 통해 93.3%의 분류 정확도를 보여주었으며, 최근에 공개된 Inception v4 module의 경우 96.9% 분류 정확도를 보여주는 우수한 신경망이다[3]. 현재 Inception v3 기반 GoogLeNet에 대한 Fine tuning 방법 및 Tensorflow 코드가 공개되어 있으며, 본 연구팀은 해당 코드를 통해 신경망 fine-tuning 과정을 수행하였다[6].

III. 실험 환경 및 결과

본 논문에서 구성된 모든 데이터베이스는 훈련 및 테스트 영상 비율을 8:2로 구성하였다. 즉, 훈련 영상의 경우 적외선 영상 데이터베이스는 400

장, 나머지 영상 데이터베이스는 5102장으로 구성 하였으며, 테스트 영상 데이터베이스는 적외선 영상 100장으로 구성하였다. 신경망 학습을 위한 Learning rate는 0.001, Iteration은 600회로 통일하였으며, 에러 역전파를 위해 확률 기울기 하강(Stochastic Gradient Descent, SGD)방법을 사용하였다.

Training \ Test		Test				
		IR	RGB	RGB + random	Gray	Gray + random
Training	IR	100%	85.9%	94.5% (+8.6)	85.2%	89.1% (+3.9)
	RGB					
Test	IR					
	RGB					

표 1. 데이터베이스 학습에 따른 적외선 영상 분류 정확도

표1은 준비한 여러 훈련용 영상 데이터베이스 학습에 따른 적외선 영상 분류 정확도를 보여준다. 실험 결과, 적외선 영상을 훈련시켰을 때 가장 좋은 성능인 100%의 분류 정확도를 보여주었으며, 컬러 영상, 회색 영상을 훈련시켰을 경우 각각 85.9%, 85.2%의 분류 정확도를 보여주었다. 적외선 영상 훈련 결과의 경우 훈련용 영상이 테스트용 영상과 같은 영상 환경에서 촬영되었다는 점에서 당연하다 할 수 있으며, 컬러·회색 영상 훈련 결과의 경우 둘 사이의 정확도 차이가 미비하여, 적외선 영상 객체 분류에 있어 객체의 색상 정보는 중요치 않다는 점을 확인할 수 있었다.

컬러 + 랜덤 패턴 영상, 회색 + 랜덤 패턴 영상을 훈련시켰을 경우 적외선 영상 분류 정확도는 각각 94.5%, 89.1%로써, 컬러 영상의 경우 +8.6%, 회색 영상의 경우 +3.9%의 분류 정확도가 향상되었음을 확인하였다. 이는 비교적 간단한 랜덤 패턴 추가를 통해서도 적외선 광원으로 인한 Speckle pattern 학습에 도움이 되었음을 의미하며, 본 연구결과와 분석을 통해 향후 정밀한 Speckle pattern 모델링을 통한 분류 성능 개선 및 적외선 영상 객체 검출 시스템에 대한 연구를 수행할 예정이다.

III. 결론

본 논문에서는 적외선 영상에 대한 객체 분류 신경망 학습을 위해 컬러 영상을 효율적으로 활용할 수 있는 방법에 대해 알아보았다. 목표 객체에 대해 컬러 영상을 확보하고, 인위적 패턴을 더하여 유사 적외선 영상을 만들고자 하였으며, 해당 영상 학습 및 적외선 영상 객체 분류 실험을 통해 단순한 컬러 영상 학습의 경우에 비해 8.6%의 정확도 향상이 있음을 확인하였다. 마지막으로 해당 실험 결과를 통해 인위적 패턴이 더해진 컬러 영상을 학습시킴으로써 효율적인 적외선 영상 분류가 가능함을 확인하였다.

ACKNOWLEDGMENT

This work was supported by ETRI R&D Program (15YI1200), funded by the government of Korea.

참고 문헌

- [1] The PASCAL Visual Object Class Homepage (<http://host.robots.ox.ac.uk/pascal/VOC/>)
- [2] IMAGENET (<http://image-net.org/>)
- [3] Christian Szegedy et al. 'Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning', arXiv, 2016
- [4] Karen Simonyan, Andrew Zisserman, 'Very Deep Convolutional Networks for Large-Scale Image Recognition', ICLR 2015
- [5] OTCBVS Benchmark Database, (<http://vcip1-okstate.org/pbvs/bench/>)
- [6] Inception v3 fine tuning instructions, (<https://github.com/tensorflow/models/tree/master/inception>)