Journal of Korea Multimedia Society Vol. 22, No. 10, October 2019(pp. 1133-1141)
https://doi.org/10.9717/kmms.2019.22.10.1133

심층 학습 모델을 이용한 수피 인식

김 민 기†

Bark Identification Using a Deep Learning Model

Min-Ki Kim†

ABSTRACT

Most of the previous studies for bark recognition have focused on the extraction of LBP-like statistical features. Deep learning approach was not well studied because of the difficulty of acquiring large volume of bark image dataset. To overcome the bark dataset problem, this study utilizes the MobileNet which was trained with the ImageNet dataset. This study proposes two approaches. One is to extract features by the pixel-wise convolution and classify the features with SVM. The other is to tune the weights of the MobileNet by flexibly freezing layers. The experimental results with two public bark datasets, BarkTex and Trunk12, show that the proposed methods are effective in bark recognition. Especially the results of the flexible tunning method outperform state-of-the-art methods. In addition, it can be applied to mobile devices because the MobileNet is compact compared to other deep learning models.

Key words: Bark Identification, MobileNet, Deep Learning Model, SVM

1. 서 론

도시화에 따른 반동으로 현대인들의 자연에 대한 관심은 점차 증대되고 있다. 그렇지만 일반인들은 주변에서 볼 수 있는 나무들의 이름조차 모르고 있다. 스마트폰의 보급과 더불어 영상 인식 기술을 활용하여 일반인들도 쉽게 나무를 식별할 수 있는 자동 인식 방법에 대한 연구가 활발히 이루어지고 있다. 나무의 식별은 과실(fruit), 꽃(flower), 잎(leaf), 수피(bark) 등에 나타난 시각적 정보를 통해 이루어진다. 본 연구는 수피 영상을 이용하여 나무를 식별하기 위한 연구로, 수피는 과실, 꽃, 잎 등에 비하여 클래스 내 유사도가 크지 않아 상대적으로 인식이 어려운 것으로 알려져 있다[1]. 그렇지만 수피는 계절에 제약을 받지 않고, 상대적으로 장기간 시각적 형태가 변하지 않는 장점이 있다[2]. 또한 수피에서 추출한 특징을 꽃이나 잎 등에서 추출한 정보와 결합하면 나무 식별에 대한 성능을 향상시킬 수 있다[1,3-5].

수피 인식을 위하여 수피에 나타난 시각적 특징을 효과적으로 추출하기 위한 다양한 연구가 진행되어 왔다. 수피의 질감(texture)을 추출하기 위한 초기 연구에서 [6]은 런 길이(run-length), 히스토그램(histogram), 동시발생행렬(co-occurrence matrix), 자기상관(auto-correlation)을 사용하였으며, [7]은 동시발생행렬을 이용하여 명도 영상에서 추출한 특징과 이진 영상에서 추출한 화소 연결길이(pixel connection length) 정보를 결합한 특징을 제안하였다. [1]은 스케일 및 회전 변형에 강인한 SIFT로 특징점을 추출하고 각 특징점 주변의 국소 특징을 방향 히스토그램으로 기술하였다. 국부적 특징을 효과적으로 나타내는 것으로 알려진 국소이진패턴(LBP) 히스토그램 방식은 특징 추출 방식이 단순하면서도 히스토그램을 이용하여 질감을 효과적으로 기술할 수 있기 때문에 수피 인식을 위한 특징 추출 방식으로

※ Corresponding Author : Min-Ki Kim, Address: (660-701) Jinju-daero 501, Jinju, Gyeongnam, Korea, TEL : +82-55-772-1393, FAX : +82-55-772-1389, E-mail : mkkim@gnu.ac.kr

Receipt date : Aug. 28, 2019, Approval date : Sep. 18, 2019
† Dept. of Computer Engineering, Gyeongsang National University Engineering Research Institute

LBP 특징을 개선하려는 꾸준한 연구가 이루어져 왔다. [8]은 중심 화소에서 반경이 R인 이웃 화소 P개에 대한 LBP를 $LBP_{P,R}$로 나타냈을 때, $LBP_{8,1}$을 $LBP_{8,r}$ $(r \in R1, R2, ..., R8)$로 확장한 다중스케일 LBP (MS-LBP)을 제안하였고, [9]는 평균 필터, LBP 필터링, Gaussian 필터, 피라미드 변환을 이용한 총 4종류의 다중블럭 LBP(Multi-Block LBP)를 제안하였다. 제안된 특징들은 기존의 LBP에 비하여 식별력이 우수한 것으로 나타났다. [10]은 나선형 패턴을 적용하여 회전 불변 특징을 추출하였고, [11]은 명도 영상에서 LBP를 구하는 방식에서 벗어나 명도 변화에 대한 크기 및 방향 정보를 영상으로 나타낸 후 여기서 LBP 특징을 추출하는 방식을 제안하였다. [2]와 [12]는 중심화소와 이웃화소의 통계적 특징을 비교하는 새로운 이진패턴을 도입하였다.

이와 같이 수피에 나타난 질감을 효과적으로 추출하기 위해 LBP 기반의 다양한 연구들이 진행되어 왔으나 수피 인식에 대한 성능을 그다지 만족스러운 결과를 나타내지 못하고 있다. 이러한 수피 인식 성능의 한계를 극복하기 위한 대안으로 딥 러닝 기술이 대두되고 있다. [13]은 심층 합성곱 신경망(Deep Convolutional Neural Network, DCNN) 모델 중 하나인 AlexNet[14]을 이용하여 특징을 추출하고 최종 단계인 완전연결신경망과 분류기를 인식 대상에 맞게 교체하여 잎 인식을 수행하였고, [15]는 ResNet 구조[16]를 갖는 ResNet18과 ResNet34를 이용하여 수피 인식을 수행하였다. [5]는 수피뿐만 아니라 과일, 꽃, 잎 등 식물의 다양한 기관에 대한 영상을 각각 분리된 DCNN에 입력시켜 특징을 추출하고 각 DCNN에서 추출된 특징을 결합하여 식물 인식(plant identification)을 인식하는 방법을 제안하였다.

본 연구에서는 딥 러닝 모델을 이용하여 수피 인식을 위한 효과적인 특징 추출 및 학습 방법을 실험을 통해 탐색하고자 한다. 일반적으로 DCNN의 앞단 레이어의 출력은 색상이나 가장자리와 같은 저차원의 특징을 식별하고, 뒤쪽 레이어로 갈수록 낮은 차원의 특징들이 결합되어 모양이나 개체들을 식별하는 것으로 알려져 있다. 본 연구는 이러한 DCNN의 특성에 착안하여 병목 특징(bottleneck feature)을 이용하는 기존 연구와는 달리 중간 레이어에서 생성되는 특징을 비교 분석하여, 그 결과를 토대로 수피 인식에 적합한 특징을 검출하였다. 본 논문의 구성은 다음과 같다. 2장에서는 합성곱 신경망(CNN)을 구성하는 기본 요소들과 다양한 딥 러닝 모델에 대하여 살펴본다. 3장에서는 본 연구에서 채택한 MobileNet 모델을 이용하여 수피 영상의 시각적 특징을 추출하고 학습하는 방법을 기술한다. 4장에서는 실험을 통해 수피 인식에 적합한 특징 추출 및 학습 방법을 탐색한다. 그리고 기존 연구결과들과 객관적인 비교가 가능하도록 관련 연구에서 널리 사용하고 있는 수피 데이터세트를 사용하여 수피 인식에 대한 실험 결과를 제시한다. 끝으로 5장에서 결론 및 향후 과제에 대하여 논의한다.

2. CNN과 딥 러닝 모델

2012년 ImageNet 대회에서 AlexNet[14]이라는 DCNN 모델이 전통적 방식의 특징 추출 방식에 의한 인식 성능을 크게 뛰어넘는 주목할 만한 성과로 우승을 차지하면서 DCNN에 대한 폭발적인 관심을 불러일으켰다. DCNN은 합성곱 레이어, 비선형화 유니트, 부샘플링 레이어의 조합으로 구성된 CNN의 기본 단위를 계층적으로 구성하여 저수준, 중간수준, 고수준의 특징 추출 능력을 갖추게 된다.

CNN의 구성 요소를 구체적으로 살펴보면 합성곱 레이어(convolution layer), 풀링 레이어(pooling layer), 활성화 함수(activation function), 배치 정규화(batch normalization), 드롭아웃(dropout), 완전연결 레이어(fully connected layer)로 나누어 볼 수 있다[17]. 합성곱 레이어에서는 커널의 크기에 대응되는 영상의 영역들로 부터 합성곱 연산을 통해 국소적으로 서로 연관된 특징을 추출한다. 식 (1)은 합성곱 연산에 의한 특징 F_n^l을 나타낸 것으로 $I_{x,y}$는 영상의 공간적 위치, k_n^l은 l번째 레이어의 n번째 커널, 그리고 *는 합성곱 연산을 나타낸 것이다.

$$F_n^l = \left(I_{x,y} * K_n^l\right) \tag{1}$$

합성곱 레이어에서 추출한 특징은 합성곱 연산이 이루어지는 영상의 위치에 따라 달라진다. 풀링 레이어에서는 이러한 위치 변화에 따른 변형을 흡수하기 위하여 합성곱 레이어에서 추출한 특징을 해당 영역을 대표하는 특징으로 변환하는 역할을 수행한다. 식 (2)를 적용하여 레이어 l의 입력특징 맵 $F_{x,y}^l$에 대하여 풀링 함수 f_p를 적용하여 레이어 l의 출력특징 맵

Z_l을 구한다.

$$Z_l = f_p\left(F_{x,y}^l\right) \tag{2}$$

활성화 함수는 입력과 각 노드의 가중치를 곱한 값들의 합에 대하여 신경망의 출력을 결정하는 함수로 시그모이드(sigmoid), 하이퍼볼릭탄젠트(tanh), 맥스아웃(maxout), ReLU(Rectified Linear Unit) 등이 있다. 적절한 활성화 함수의 선택은 학습 과정을 촉진한다. 배치 정규화는 활성화 함수의 활성화 값 또는 출력 값을 정규화하는 것으로, 식 (3)을 이용하여 데이터 분포가 평균이 0, 분산이 1이 되도록 정규화 한다. 식 (3)에서 N_n^l은 정규화된 특징 맵, T_n^l는 입력 특징 맵, σ는 특징 맵의 분산을 나타낸다. 이를 통해 기울기 소실(vanishing gradient) 문제와 과적합(overfitting) 문제를 줄일 수 있다.

$$N_n^l = \frac{T_n^l}{\sigma^2 + \sum_i T_i^l} \tag{3}$$

드롭아웃은 학습 단계마다 확률적으로 일부 뉴런을 제거함으로써 어떤 특징이 특정 뉴런에 고정되는 것을 막아 가중치의 균형을 잡도록 하는 것이다. 이를 통해 과적합을 방지한다. 완전연결 레이어는 분류를 목적으로 망의 끝에서 사용된다. 이 레이어는 이전 레이어의 입력을 출력 레이어 전체에 연결한다.

CNN의 기본 구조를 적층하여 저수준, 중간수준, 고수준의 특징 추출이 가능하다. CNN의 깊이와 너비가 늘어날수록 복잡한 학습 문제를 해결할 수 있을 것으로 기대되나, 학습을 위한 파라미터의 개수와 연산 양이 늘어나며 이에 따른 과적합과 기울기 소실 문제를 해결해야 한다. AlexNet으로부터 촉발된 딥러닝 모델은 VGGNet, GoogleNet, ResNet, Inception ResNet, DenseNet, MobileNet 등 다양한 형태의 DCNN 모델들로 확장되었다[18]. AlexNet은 11×11, 5×5, 3×3의 합성곱 층을 포함하여 7개 레이어로 구성된 반면, InceptionResNet은 572개의 레이어를 갖는 매우 깊은 신경망이다. VGGNet[19]은 16개의 레이어를 갖는 신경망으로, 상대적으로 넓은 영역에 나타나는 특징을 추출하기 위하여 AlexNet은 11×11 필터를 사용한데 비하여, VGGNet은 3×3 필터를 연속적으로 적용하였다. GoogleNet은 큰 신경망을 희소(sparse)하게 연결한 구조인 Inception 모듈을 도입하여 학습시킬 파라미터의 양을 줄임으로써 신경망을 더욱 깊게 만들 수 있게 하였다. ResNet은 기울기 소실 문제를 해결하기 위하여 바로가기 연결(short-cut connection)을 추가하였다. InceptioResNet은 GoogleNet에서 제안된 Inception 모듈과 ResNet의 Residual 연결을 조합한 구조이며, DenseNet은 상위 레이어의 출력을 나머지 하위 레이어로 연결하여 기존에 학습된 특징이 재사용되는 구조이다[17]. MobileNet은 입력에 대한 필터링과 추출된 특징의 결합을 통한 출력을 한 단계의 합성곱으로 수행하는데 반해, 이것을 각각 필터링 레이어와 결합 레이어로 분리하여 연산 시간과 모델의 크기를 대폭 줄였다[20]. Table 1은 딥 러닝 모델들의 크기와 파라미터 개수를 비교한 것으로, MobileNet은 모델 크기와 파라미터 개수가 상대적으로 매우 작아 모바일 기기에 적용할 수 있는 장점이 있다.

3. MobileNet 모델을 이용한 수피 인식

본 연구에서는 수피 인식을 위하여 모바일 기기에 적용이 가능한 MobileNet을 이용하였다. ImageNet [21] 데이터세트에 대한 실험 결과 MobileNet은 GooglelNet에 비하여 성능이 우수하고, VGG16 모델에 버금가는 성능을 나타내면서도 대략적으로 모델의 크기가 32배나 작고 계산시간은 27배 빠른 장점이 있기 때문이다. 본 장에서는 먼저 MobileNet 모델의 구조를 살펴본 후 전이 학습(transfer learning)과 미세 조정(fine tunning)을 통해 MobileNet을 학습시켜 수피 인식을 수행하는 방법에 대하여 기술한다.

3.1 MobileNet 모델의 구조

MobileNet 모델[20]은 깊이단위(depth-wise)의 분리 가능한 합성곱에 기반하고 있다. 깊이단위의 합성곱은 각 입력 채널에 대하여 단일 필터를 적용하

Table 1. Comparison of deep learning models

Model	Model Size	No. of Parameters
VGG16	528MB	138,357,544
Resnet50	98MB	25,636,712
InceptionV3	92MB	23,851,784
InceptionResNetV2	215MB	55,873,736
MobileNet	16MB	4,253,864

고, 그 출력을 결합하기 위하여 화소단위(pixel-wise) 합성곱을 적용한다. 즉, 표준 합성곱은 입력에 대한 필터링과 추출된 특징의 결합을 통한 출력을 한 단계에서 수행하지만, MobileNet에서는 이것을 각각 필터링 레이어와 결합 레이어로 분리하여 수행한다. 이것은 연산 시간과 모델의 크기를 획기적으로 줄이는 효과를 갖는다. MobileNet은 필터링과 결합, 두 레이어에서 모두 배치 정규화와 ReLU 연산을 사용한다. MobileNet의 전체적인 구조는 Table 2와 같다.

3.2 수피 인식을 위한 MobileNet 학습

심층 합성곱 신경망(DCNN)을 학습시키기 위해서는 대용량의 데이터세트가 필요할 뿐만 아니라 고성능의 컴퓨터를 사용하더라도 매우 긴 학습 시간이 요구된다. ImageNet[21] 데이터세트는 약 3백만 개의 영상들로 구성되어 있다. 이에 비하여 수피 인식 연구에서 널리 사용되어온 BarkTex, AFF, Trunk12 데이터세트의 크기는 수백 개에 불과하다. 이러한 경우 비록 학습에 사용된 데이터가 목적 데이터(target data)가 아니더라도 대용량 데이터로 훈련된 기존의 DCNN을 이용하면 전이 학습(transfer learning)이나 미세 조정(fine tuning)을 통해 신경망을 효과적으로 학습시킬 수 있다[22, 23].

Table 2. MobileNet Architecture [20]

Type / Stride		Filter Shape	Input Size
Conv /s2		3×3×3×32	224×224×3
Conv dw/s1		3×3×32 dw	112×112×32
Conv /s1		1×1×32×64	112×112×32
Conv dw/s2		3×3×64 dw	112×112×64
Conv /s1		1×1×64×128	56×56×64
Conv dw/s1		3×3×128 dw	56×56×128
Conv /s1		1×1×128×128	56×56×128
Conv dw/s2		3×3×128 dw	56×56×128
Conv /s1		1×1×128×256	28×28×128
Conv dw/s1		3×3×256 dw	28×28×256
Conv /s1		1×1×256×256	28×28×256
Conv dw/s2		3×3×256 dw	28×28×256
Conv /s1		1×1×256×512	14×14×256
5 ×	Conv dw/s1 Conv/s1	3×3×512 dw 1×1×512×512	14×14×512 14×14×512
Conv dw/s2		3×3×512 dw	14×14×512
Conv /s1		1×1×512×1024	7×7×512
Conv dw/s2		3×3×1024 dw	7×7×1024
Conv /s1		1×1×1024×1024	7×7×1024
Avg Pool /s1		Pool 7×7	7×7×1024
FC /s1		1024×1000	1×1×1024
Softmax /s1		Classifier	1×1×1000

본 연구에서는 MobileNet을 이용한 특징 추출과 탄력적 미세 조정을 통해 수피 인식을 수행하는 두 가지 접근을 시도하였다. 첫 번째 방식은 Fig. 1에 제시된 바와 같이 사전 학습된 파라미터를 통해 추출한 특징을 주성분분석(Principal Component Analysis, PCA)을 통해 차원을 축소시킨 후 이들로부터 수피를 식별하기 위하여 지지벡터머신(Support Vector Machine, SVM)을 사용한다. 기존 연구들[5,13,15]은 병목 특징으로 수피 인식을 수행하는데 비하여, 본 연구에서는 그 이전의 레이어들에서 나타나는 특징들을 이용하여 수피 인식을 수행한다. 일반적으로 DCNN의 앞쪽에 나타나는 합성곱 특징은 색상이나 에지와 같은 낮은 차원의 특징을 식별하고, 뒤쪽에서는 이런 낮은 차원의 특징을 결합한 모양이나 개체들을 식별한다. 따라서 중간 이후 계층이 수피 영상의 중요한 특징을 나타낸다고 볼 수 있다.

두 번째 방식은 완전연결신경망 이전의 모든 레이어를 동결시켜 미세 조정하는 기존의 방식에서 벗어나 동결시키는 레이어의 범위를 탄력적으로 조절한다. ImageNet 데이터세트로 학습된 MobilleNet은 1,000개의 클래스를 식별하도록 학습되었기 때문에 특정 물체의 특징에 종속되지 않는 일반적인 특징을 추출한다고 볼 수 있다. 따라서 수피 인식에 적합한 특징을 추출할 수 있도록 수피 영상 데이터세트를 이용한 학습을 통해 신경망의 파라미터를 조정하여야 한다. 우선 MobileNet의 병목 특징을 입력으로

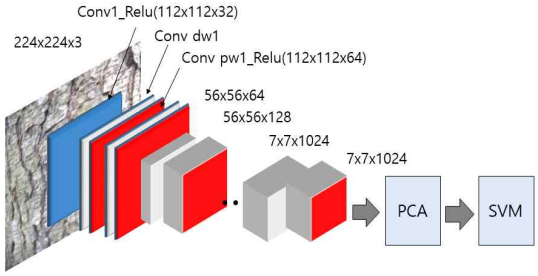

Fig. 1. Overview of the transfer learning using MobileNet.

Table 3. Architecture of the newly added back-end of the MobileNet

Input	7 × 7 × 1024
Pooling	Avg. Pool_size: 7 × 7
Flatten	
Fully connected	units: 128, activation: ReLU
Dropout	rate: 0.5
Fully connected	units: n_classes, activation: softmax

받아 수피 인식을 수행하기 위한 얕은 신경망을 추가하였다. 새로 추가한 신경망은 Table 3에 제시된 바와 같이 1개의 풀링 레이어와 2개의 완전연결 레이어로 이루어진 매우 단순한 구조이다. 활성화 함수는 ReLU를 사용하였고, 최종 단계에서는 softmax를 적용하였다. 그리고 신경망의 과적합 문제를 해소하기 위하여 드롭아웃 비율은 0.5로 설정하였다.

4. 실험 결과 및 분석

4.1 실험 데이터

MobileNet을 이용한 수피 인식 방법에 대한 타당성을 확인하고, 그 결과를 기존 연구들과 비교하기 위하여 수피 인식에 널리 사용되어 왔던 BarkTex, Trunk12 데이터세트를 실험 데이터로 사용하였다. BarkTex 데이터세트는 총 408개의 컬러 영상으로 구성되어 있으며, 영상의 해상도는 256×384 이다. 6개의 클래스로 구성되어 있으며, 각 클래스는 68개 컬러 영상으로 이루어져 있다. Fig. 2에 제시된 바와 같이 나무의 몸통(trunk)이 영상의 중앙에 위치하고 있지만 몸통의 크기에 따라 배경이 포함되기도 한다. 얼핏 보면 서로 다른 종류의 나무로 보이지만 실은 모두 한 종류의 나무에서 획득한 영상들이다. Trunk12 데이터세트는 총 393개의 고해상도(3,000×4,000) 컬러 영상으로 이루어진 12개의 클래스로 구성되어 있다. 각 클래스는 최소 30개에서 최대 45개의 영상으로 이루어져 있다. 실험에 사용한 데이터들은 분류할 클래스의 개수는 많지 않지만 Fig. 2에 나타난 바와 같이 동일한 클래스 내의 영상들의 변형이 크게 나타나고, Fig. 3에 제시된 바와 같이 클래스 간 영상들 간의 유사도가 커서 분류하는데 어려움을 내포하고 있다.

Fig. 2. Sample images selected at BarkTex.

4.2 MobileNet과 SVM을 이용한 수피 인식

MobileNet은 총 86개의 레이어로 구성되어 있으며, 화소단위의 합성곱을 중심으로 살펴보면 13개 그룹으로 나누어 볼 수 있다. 어떤 합성곱 특징이 수피를 식별하는데 보다 효과적인지 파악하기 위하여 병목 특징은 물론 그 이전 레이어들의 합성곱 특징을 추출하여 수피 인식 실험을 수행하였다. Fig. 4는 각 레이어별로 화소단위 합성곱 특징을 추출하여 PCA에 의하여 차원을 100으로 축소한 후 SVM으로 식별한 인식 결과를 나타낸 것이다. SVM 분류기를 학습시키기 위하여 전체 데이터의 80%를 훈련 데이터로 사용하고, 나머지 20%를 평가 데이터로 사용하였다.

Fig. 3. Sample images selected at Trunk12.

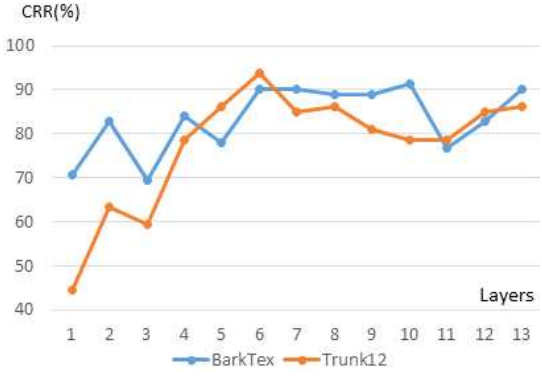

Fig. 4. SVM classification results by the features of each convolution layer.

수피에 대한 식별 성능은 식 (4)에 제시된 정인식률(Correct Recognition Rate, CRR)로 측정하였다. N은 전체 테스트 영상의 개수, C는 올바르게 식별한 영상의 개수를 나타낸다. 6번째와 13번째의 합성곱 특징이 수피 인식에 대한 식별력이 높은 것으로 나타났으며, 6번째 합성곱 특징을 사용하였을 때 Bark Tex는 90.24%, Trunk12는 93.67%로 가장 높은 수피 인식 성능을 보였다.

$$CRR(\%) = \frac{C}{N} \times 100 \quad (4)$$

4.3 탄력적 미세 조정에 의한 수피 인식

일반적으로 사전 학습된 DCNN을 미세 조정할 때 완전연결신경망 이전의 모든 레이어를 동결한 후 새로 추가된 얕은 신경망에 대하여 학습을 수행하는데, 본 연구에서는 동결되는 레이어를 탄력적으로 조정하였다. ImageNet 데이터세트로 사전 학습된 Mobile Net은 범용적인 특징에 대한 파라미터 학습이 이루어진 상태이므로, 얕은 신경망 앞단의 존재하는 이들 파라미터들도 학습 과정에서 수피 인식에 적합한 파라미터로 조정하는 것이 인식 성능에 영향을 미칠 수 있기 때문이다.

미세 조정을 위한 학습 과정도 SVM 분류기를 학습시킬 때와 동일하게 전체 데이터의 80%를 훈련 데이터로 사용하고, 나머지 20%를 평가 데이터로 사용하였다. 미세 조정 과정에서 MobileNet 레이어의 동결은 처음부터 N번째 화소단위 합성곱 레이어까지 동결하는 방식으로 학습을 진행시켜 수피 인식 실험을 수행한 결과 Fig. 5의 결과를 얻었다. N이 작을수록 조정되는 파라미터의 개수는 늘어나는데, 이것은 수피 훈련 데이터의 학습에 의하여 영향을 받는 파라미터가 증가함을 의미한다. 이것은 신경망의 파라미터가 수피 인식에 적합하게 튜닝되는 바람직한 측면이지만, 다른 한편으로는 훈련데이터가 충분하지 않을 경우 훈련 데이터에 과적합되는 단점도 내포한다.

BarkTex와 Trunk12는 각각 10번째와 8번째 합성곱 레이어까지 동결시켰을 때 가장 좋은 인식 성능을 보였으며, 8번째와 13번째의 합성곱 레이어까지 동결시켜 미세 조정을 하였을 때 BarkTex와 Trunk12 두 데이터세트 모두 상대적으로 높은 식별력을 보여주었다. 두 데이터세트를 모두 고려할 때, 8번째 합성곱 레이어까지 동결시키고 그 이후 레이어들을 학습시켰을 때 BarkTex는 98.78%, Trunk12는 92.41%로 가장 좋은 인식 성능을 보였다.

Fig. 5. MobileNet classification results by flexible frozen layers.

4.4 수피 인식 성능 비교

제안된 방법의 수피 인식에 대한 효과성을 비교하기 위하여 동일한 실험 데이터를 이용한 최근의 연구 결과들을 정리하여 Table 4에 제시하였다. BarkTex와 Trunk12는 인터넷에 공개되어 있으나, AFF는 비공개 데이터세트로 데이터를 확보하지 못하여 이번 연구에는 포함하지 못하였다.

2DSCAR[10]은 나선형 패턴을 적용하여 회전 불변 특징을 추출한 연구이고, GLBP[11]는 명도 영상뿐만 아니라 그레디언트(gradient) 영상을 이용하여 LBP를 공간적으로 확장한 연구이다. [12]와 [2]는 모두 국소 영역에 나타나는 통계적 특징을 이용한다는 공통점이 있는데, SMBP[12]는 입력 영상으로부터 통계(평균, 중앙, 최소, 최대, 분산) 영상을 만들어 각

Table 4. Performance comparison with different researches

Research \ Dataset	BarkTex	Trunk12	AFF
2DSCAR [10]	91.7%	92.9%	83.6%
GLBP [11]	94.39%	78.39%	72.21%
SMBP [12]	84.31%	71.00%	71.74%
LCoLBP/H30 [2]	92.4%	84.2%	80.7%
MobileNet6(SVM)	90.24%	**93.67%**	-
MobileNet8(FT)	**98.78%**	92.41%	-

영상을 매크로 이진패턴(MBP)으로 인코딩한 후 이들을 연결하여 국소 특징을 기술하였고, LCoLBP/H30[2]는 통계적 LBP 특징과 30개의 빈(bin)을 갖는 색상 히스토그램(hue)으로 특징을 기술하였다. 마직막 2개는 본 연구에서 제안한 방식으로 MobileNet6(SVM)은 MobileNet에서 6번째 화소단위 합성곱 레이어에서 추출된 특징을 SVM 분류기를 이용하여 수피 인식을 수행하는 방식을 나타내며, MobileNet8(FT)는 8번째 화소단위 합성곱 레이어까지 동결시킨후 미세 조정을 통해 수피 인식을 수행하는 방식을 나타낸다. MobileNet6(SVM)은 Trunk12에 대하여 가장 우수한 성능을 보였으나, BarkTex에는 상대적으로 낮은 성능을 보였으며, MibileNet8(FT)은 Bark Tex에서 가장 우수한 성능을 나타내면서도 Trunk12에 대해서도 만족할만한 성능을 나타내었다.

5. 결 론

영상 인식 기술을 적용하여 나무를 자동으로 식별할 때, 수피는 꽃이나 잎에 비하여 계절적 제약이 적고 상대적으로 긴 시간동안 시각적 특징이 변하지 않는 장점이 있다. 기존 연구들은 LBP 기반의 국소 특징을 효과적으로 추출하기 위한 방법을 찾는데 집중하여 왔으나, 수피 인식 성능은 크게 향상되지 않았다. 그리고 딥 러닝 기술을 적용하여 수피 인식을 수행한 연구는 아직까지 저조한 상태이다. 이것은 심층신경망을 훈련시키는데 필요한 충분한 양의 수피 영상 데이터세트를 구하기 어려운 현실적 측면이 있기 때문인 것으로 판단된다. 본 연구에서는 이러한 문제를 해결하기 위하여 1,000개의 클래스, 약 3백만개의 영상들로 학습된 MobileNet을 이용하여 수피의 특징을 추출하고 효과적으로 학습시키기 위한 방법을 실험을 통해 탐색하였다. MobileNet의 화소단위 합성곱에서 추출한 특징을 SVM으로 분류하는 방식과 MobileNet을 탄력적으로 동결시켜 미세 조정하는 두 가지 방식을 제시하였다. BarkTex와 Trunk12 데이터세트를 이용하여 실험한 결과 두 방식 모두 기존 연구에 비해 만족할만한 수피 인식 성능을 보여 제안한 방식이 수피 인식에 효과적임을 확인할 수 있었다. 특히 수피 인식에 대한 응용을 고려할 때 MobileNet을 이용함으로써 모바일 기기에 적용이 가능한 장점이 있다. 향후 연구에서는 수피뿐만 아니라 꽃과 잎에 대한 정보를 함께 활용하여 나무 인식 성능을 개선할 수 있는 방법을 탐색할 계획이다.

REFERENCE

[1] S. Fiel and R. Sablatnig, “Automated Identification of Tree Species from Images of the Bark, Leaves and Needles,” *Proceedings of the Computer Vision Winter Workshop,* pp. 67-74, 2011.

[2] R. Ratajczak, S. Bertrand, C. Crispim-Junior, and L. Tougne, "Efficient Bark Recognition in the Wild," *Proceedings of the International Conference on Computer Vision Theory and Applications*, pp. 240-248, 2019,

[3] S. Kim, B. Kim, and D. Kim, "Tree Recognition for Landscape Using by Combination of Features of its Leaf, Flower, and Bark," *Proceedings of the SICE Annual Conference,* pp. 1147-1151, 2011.

[4] S. Bertrand, G. Cerutti, and L. Tougne, "Bark Recognition to Improve Leaf-based Classification in Didactic Tree Species Identification," *Proceedings of the International Conference on Computer Vision Theory and Applications,* pp. 435-442, 2017.

[5] A. He and X. Tian, “Multi-Organ Plant Identification with Multi-Column Deep Convolutional Neural Networks,” *Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics,* pp. 20-25, 2016.

[6] Y.-Y. Wan, J.-X. Du, D.-S. Huang, Z. Chi, Y.-M. Cheung, X.-F. Wang, and G-.J. Zhang, “Bark Texture Feature Extraction Based on Statistical Texture Analysis,” *Proceedings of the International Symposium on Intelligent Multimedia, Video and Speech Processing,* pp. 482-485, 2004.

[7] J. Song, Z. Chi, J. Liu, and H. Fu, “Bark Classification by Combining Grayscale and Binary Texture Features,” *Proceedings of the International Symposium on Intelligent Multimedia, Video and Speech Processing,* pp. 450-

453, 2004.

[8] M. Sule and J. Matas, "Kernel-mapped Histograms of Multi-scale LBPs for Tree Bark Recognition," *Proceedings of the IEEE International Conference on Image and Vision Computing New Zealand,* pp. 82-87, 2013.

[9] S. Boudra, I. Yahiaoui, and A. Behloul, "A Comparison of Multi-Scale Local Binary Pattern Variants for Bark Image Retrieval," *Proceedings of the International Conference on Advanced Concepts for Intelligent Vision Systems,* pp. 764-775, 2015.

[10] V. Remes and M. Haindl, "Rotationally Invariant Bark Recognition," *Proceedings of the Joint IAPR International Workshops on Statistical Techniques in Pattern Recognition and Structural and Syntactic Pattern Recognition,* pp. 22-31, 2018.

[11] T. Le-Viet and V.T. Hoang, "Local Binary Pattern Based on Image Gradient for Bark Image Classification," *Proceedings of the International Conference on Signal Processing Systems,* Vol. 11071, 2019.

[12] S. Boudra, I. Yahiaoui, and A. Behloul, "Plant Identification form Bark: A Texture Description based on Statistical Macro Binary Pattern," *Proceedings of the International Conference on Pattern Recognition,* pp. 1530-1535, 2018.

[13] S.H. Lee, C.S. Chan, P. Wilkin, and P. Remagnino, "Deep-Plant: Plant Identification with Convolutionl Neural Networks," *Proceedings of the IEEE International Conference on Image Processing*, pp. 452-456, 2015.

[14] A. Krizhevsky, I. Sutskever, and G.E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," *Proceedings of International Conference on Neural Information Processing System,* Vol. 1, pp. 1097-1105, 2012.

[15] M. Carpentier, P. Giguere, and J. Gaudreault, "Tree Species Identification from Bark Images Using Convolutional Neural Networks," *Proceedings of the IEEE International Conference on Intelligent Robots and Systems*, pp. 1075-1081, 2018.

[16] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition,* pp. 770-778, 2016.

[17] A. Khan, A. Sohail, U. Zahoora, and A.S. Qureshi, "A Survey of the Recent Architecture of Deep Convolutional Neural Networks," *arXiv Preprint arXiv:1901.06032*, 2019.

[18] H. Rahul and R.L. Jyothi, "Convolutional Neural Networks: A Comprehensive Survey," *International Journal of Applied Engineering Research,* Vol. 14, No. 3, pp. 780-789, 2019.

[19] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," *Proceedings of International Conference on Learning Representations,* pp. 1-14, 2014.

[20] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, "MobileNets: Efficient Convolutional Netural Networks for Mobile Vision Applications," *arXiv Preprint arXiv:1704.04861*, 2017.

[21] D. Jia, D. Wei, S. Richard, L-J. Li, K. Li, and F-F. Li, "ImageNet: A Large-Scale Hierarchical Image Database," *Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition*, pp. 248-255, 2009.

[22] K. Weiss, T.M. Khoshgoftaar, and D. Wang, "A Survey on Transfer Learning," *Journal of Big Data*, Vol. 3, No. 9, pp. 1-40, 2014.

[23] M. Kim, "Contactless Palmprint Identification Using the Pretrained VGGNet Model," *Journal of Korea Multimedia Society*, Vol. 21, No. 12, pp. 1439-1447, 2018.

김 민 기

1989년 2월 중앙대학교 전자계산학과 이학사
1994년 8월 중앙대학교 전자계산학과 이학석사
1998년 2월 중앙대학교 컴퓨터공학과 공학박사

1998년~1999년 KRIC 선임연구원
1999년~2000년 KERIS 책임연구원, 연구위원
2000년~2010년 경상대학교 컴퓨터교육과 부교수
2011년~현재 경상대학교 컴퓨터과학과 교수
2013년~2014년 미국 University of Notre Dame 방문학자
2014년~현재 경상대학교 공학연구원(ERI) 자동화·컴퓨터연구센터 책임연구원
관심분야: 딥 러닝, 생체인식, 패턴인식, 영상처리