

Multi-label Classification & EfficientNet



AI 빅데이터융합 경영학과 20 김상욱



목차



주제 소개



선정 논문



논문 리뷰



앞으로의 방향성

주제 소개

돌봄 절실한데 진료 후순위...의료
취약층 :

전체 인구 중 디지털 헬스 정보에 어려움이 있는 인구는 연구에 따라 22~58%로 추정된다. 디지털 헬스
▲연령이 높을수록 ▲교육이 낮을수록 ▲경제 형편이 낮을수록 ▲직업이 없는 경우 ▲도시에서 먼 경
지가 낮은 경우 ▲편협 혼란이나 경험에 적은 경우 ▲자기 효능감이 낮을수록 부정적으로 나타났다.

주복지재단 대담회 - 의료 불균형: 필수의료 & 취약계층
"공공의료는 취약계층 진료만 !



조기발견이 중요한 '눈'
의료 취약층 -> 진료 후순위
의료불균형 -> 의료 취약층 피해



의료 모델을 통해
간이 검진 서비스
+
약 자판기 (처방)
+
심각할 경우 인근 병원 매칭(조기 진단 후)

선정 논문



안저 이미지 Multi-label 분류에 대한 논문

Multi-Label Classification of Fundus Images With EfficientNet

JING WANG¹, LIU YANG¹, ZHANQIANG HUO¹, WEIFENG HE², AND JUNWEI LUO¹

¹College of Computer Science and Technology, Henan Polytechnic University, Jiaozuo 454003, China

²Henan Polytechnic University Hospital, Jiaozuo 454003, China

Corresponding author: Jing Wang (wjasmine@hpu.edu.cn)



EfficientNet에 대한 논문

EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks

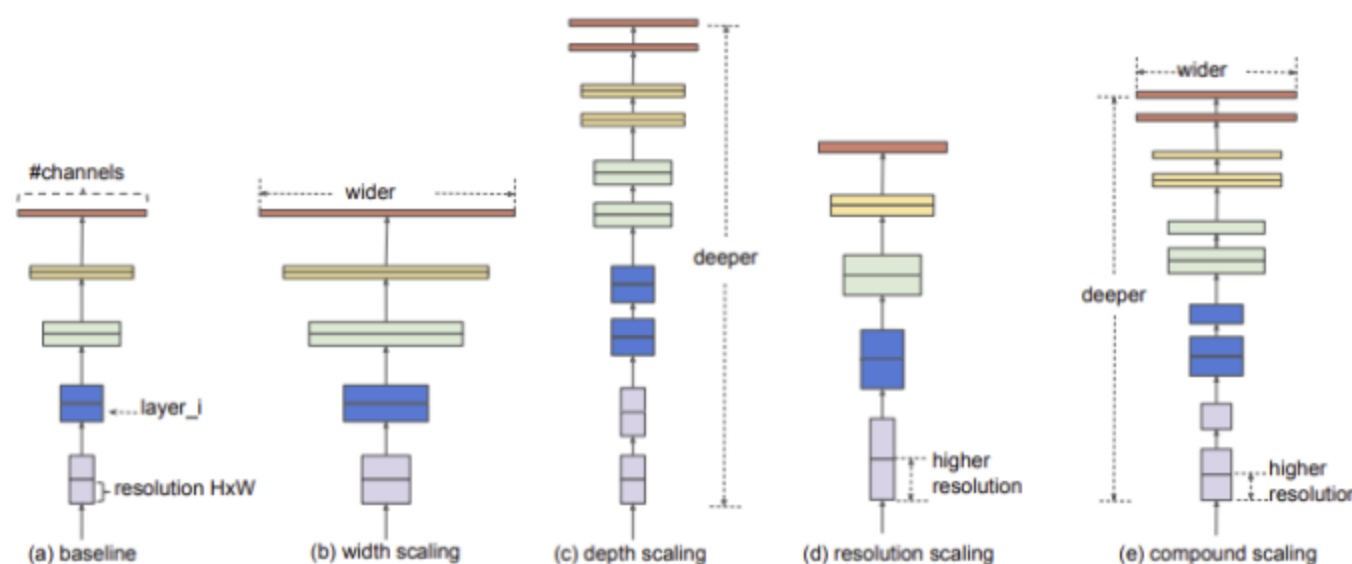
Mingxing Tan¹ Quoc V. Le¹



EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks



논문 리뷰(1)



Better accuracy? Better Efficiency?
accuracy가 어느정도 증가 → Efficiency를 증가하는 방향으로

이전 연구 : depth, width, image resolution 중에 하나만 scale up
본 연구 : 3가지를 다 고려해서 scale up (Compound scaling)

$$\begin{aligned} \text{depth: } d &= \alpha^\phi \\ \text{width: } w &= \beta^\phi \\ \text{resolution: } r &= \gamma^\phi \\ \text{s.t. } \alpha \cdot \beta^2 \cdot \gamma^2 &\approx 2 \\ \alpha \geq 1, \beta \geq 1, \gamma &\geq 1 \end{aligned}$$



$$\begin{aligned} \max_{d,w,r} \quad & \text{Accuracy}(\mathcal{N}(d, w, r)) \\ \text{s.t.} \quad & \mathcal{N}(d, w, r) = \bigodot_{i=1 \dots s} \hat{\mathcal{F}}_i^{d \cdot \hat{L}_i} (X_{\langle r \cdot \hat{H}_i, r \cdot \hat{W}_i, w \cdot \hat{C}_i \rangle}) \\ & \text{Memory}(\mathcal{N}) \leq \text{target_memory} \\ & \text{FLOPS}(\mathcal{N}) \leq \text{target_flops} \end{aligned}$$

F값을 고정하고 d,w,r을 변화하면서 max Accuracy값 구함

논문 리뷰(1)

Stage i	Operator \mathcal{F}_i	Resolution $H_i \times W_i$	#Channels C_i	#Layers L_i
1	Conv3x3	224×224	32	1
2	MBConv1, k3x3	112×112	16	1
3	MBConv6, k3x3	112×112	24	2
4	MBConv6, k5x5	56×56	40	2
5	MBConv6, k3x3	28×28	80	3
6	MBConv6, k5x5	14×14	112	3
7	MBConv6, k5x5	14×14	192	4
8	MBConv6, k3x3	7×7	320	1
9	Conv1x1 & Pooling & FC	7×7	1280	1



MBConv ?

building block의 한 종류

입력된 channel 확장 → filter를 통해 공간 정보 처리 → channel 축소

=> 연산량과 parameter 수가 적음

=> 확장-축소를 통해 표현력 극대화

problem → 큰 모델에서 값을 찾는 비용 증가

step 1

ϕ 값을 1로 고정한 뒤, 2배 더 많은 resource 사용 가능하다고 가정

→ 소규모 grid search를 통해 α, β, γ 값 구함

step 2

α, β, γ 값 고정, ϕ 값을 변화시키면서 scale up 진행



EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks



논문 리뷰(1)

Model	FLOPS	Top-1 Acc.
Baseline MobileNetV1 (Howard et al., 2017)	0.6B	70.6%
Scale MobileNetV1 by width ($w=2$)	2.2B	74.2%
Scale MobileNetV1 by resolution ($r=2$)	2.2B	72.7%
compound scale ($d=1.4, w=1.2, r=1.3$)	2.3B	75.6%
Baseline MobileNetV2 (Sandler et al., 2018)	0.3B	72.0%
Scale MobileNetV2 by depth ($d=4$)	1.2B	76.8%
Scale MobileNetV2 by width ($w=2$)	1.1B	76.4%
Scale MobileNetV2 by resolution ($r=2$)	1.2B	74.8%
MobileNetV2 compound scale	1.3B	77.4%
Baseline ResNet-50 (He et al., 2016)	4.1B	76.0%
Scale ResNet-50 by depth ($d=4$)	16.2B	78.1%
Scale ResNet-50 by width ($w=2$)	14.7B	77.7%
Scale ResNet-50 by resolution ($r=2$)	16.4B	77.5%
ResNet-50 compound scale	16.7B	78.8%

	Comparison to best public-available results						Comparison to best reported results					
	Model	Acc.	#Param	Our Model	Acc.	#Param(ratio)	Model	Acc.	#Param	Our Model	Acc.	#Param(ratio)
CIFAR-10	NASNet-A	98.0%	85M	EfficientNet-B0	98.1%	4M (21x)	[†] Gpipe	99.0%	556M	EfficientNet-B7	98.9%	64M (8.7x)
CIFAR-100	NASNet-A	87.5%	85M	EfficientNet-B0	88.1%	4M (21x)	Gpipe	91.3%	556M	EfficientNet-B7	91.7%	64M (8.7x)
Birdsnap	Inception-v4	81.8%	41M	EfficientNet-B5	82.0%	28M (1.5x)	Gpipe	83.6%	556M	EfficientNet-B7	84.3%	64M (8.7x)
Stanford Cars	Inception-v4	93.4%	41M	EfficientNet-B3	93.6%	10M (4.1x)	[‡] DAT	94.8%	-	EfficientNet-B7	94.7%	-
Flowers	Inception-v4	98.5%	41M	EfficientNet-B5	98.5%	28M (1.5x)	DAT	97.7%	-	EfficientNet-B7	98.8%	-
FGVC Aircraft	Inception-v4	90.9%	41M	EfficientNet-B3	90.7%	10M (4.1x)	DAT	92.9%	-	EfficientNet-B7	92.9%	-
Oxford-IIIT Pets	ResNet-152	94.5%	58M	EfficientNet-B4	94.8%	17M (5.6x)	Gpipe	95.9%	556M	EfficientNet-B6	95.4%	41M (14x)
Food-101	Inception-v4	90.8%	41M	EfficientNet-B4	91.5%	17M (2.4x)	Gpipe	93.0%	556M	EfficientNet-B7	93.0%	64M (8.7x)
Geo-Mean	(4.7x)						(9.6x)					

compound scale은 EfficientNet 뿐만 아니라 다른 모델에서도 효과적

더 적은 parameter 수, FLOPS에서 좋은 성능을 보임

다양한 데이터셋에서도 대부분 좋은 성능을 보임

논문 리뷰(2)

이전 연구 : 특정 안저 질환의 탐지에 대해 focus

본 연구 : 다중 레이블 분류 앙상블 모델을 제안 → 하나 이상의 안저 질환을 감지해보자!

why? 다중 레이블 분류를 사용

1. 실제 세계의 안저 이미지는 여러 안저 질환을 포함할 가능성이 높음
2. 충분한 실제 이미지를 얻기 어려움 (희귀 질병에 경우 더 심함)
3. 제한된 이미지 데이터와 노이즈 상황에서 단일 모델을 훈련하여 높은 질병 탐지 정확도 얻기 어려움

=> 전이 학습과 앙상블(ensemble)

전이 학습?

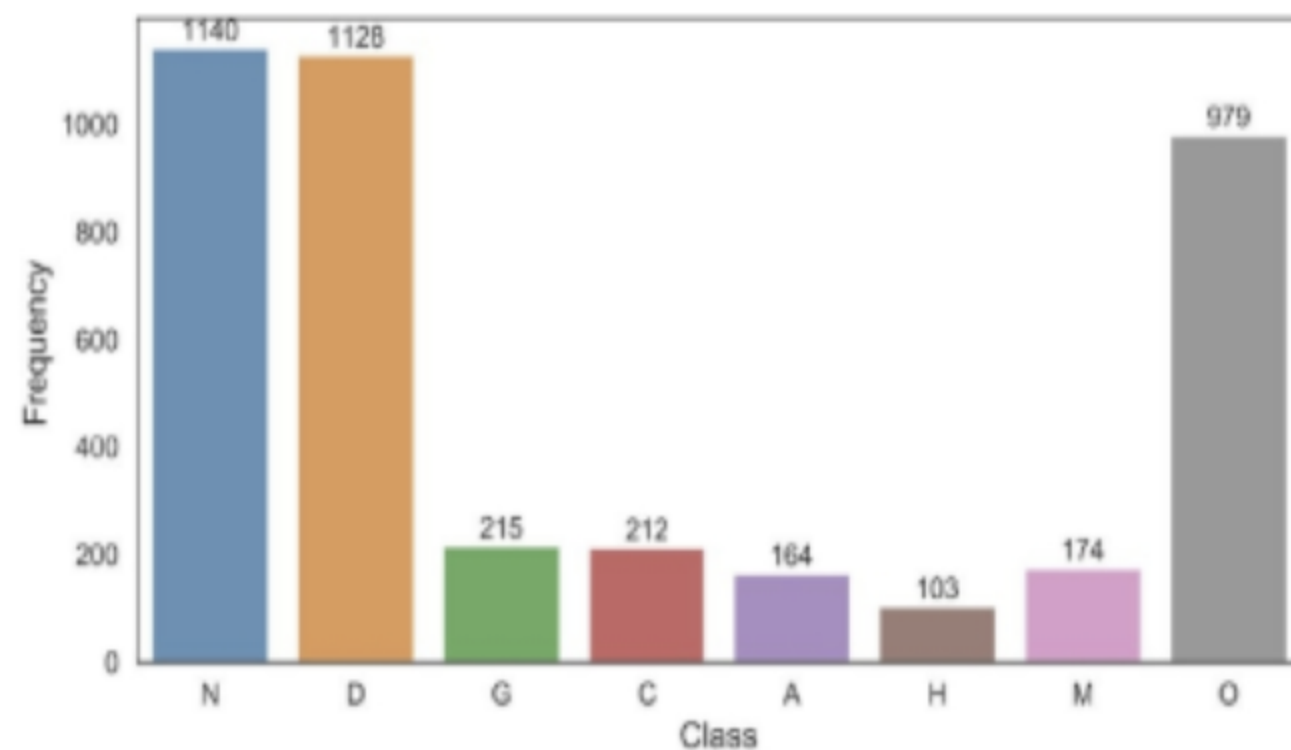
=> 작업(Task)에서 학습한 모델의 지식을 다른 관련 작업에 적용하는 것을 의미
ex) ImageNet에서 EfficientNet을 학습시키고 안저 이미지에서도 그대로 적용



Multi-Label Classification of Fundus Images With EfficientNet

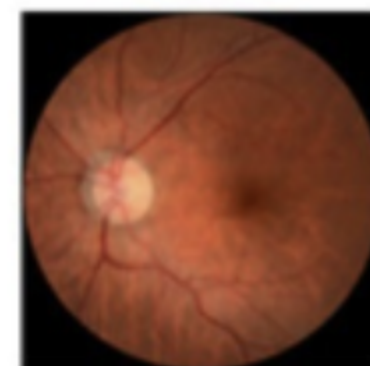


논문 리뷰(2)

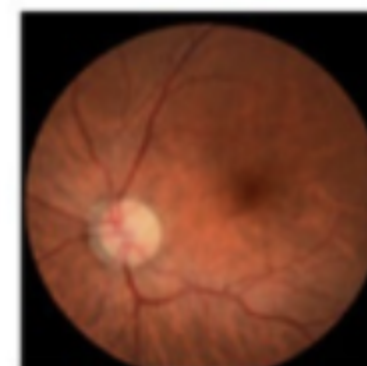


Dataset

중국의 여러 병원 기관에서 얻은 '실제' 데이터
실제 불균형 존재 -> 가중치 손실 함수



Original



Rotate



Rotate

train: test = 9:1

input image = 229x229

DataArgumentation

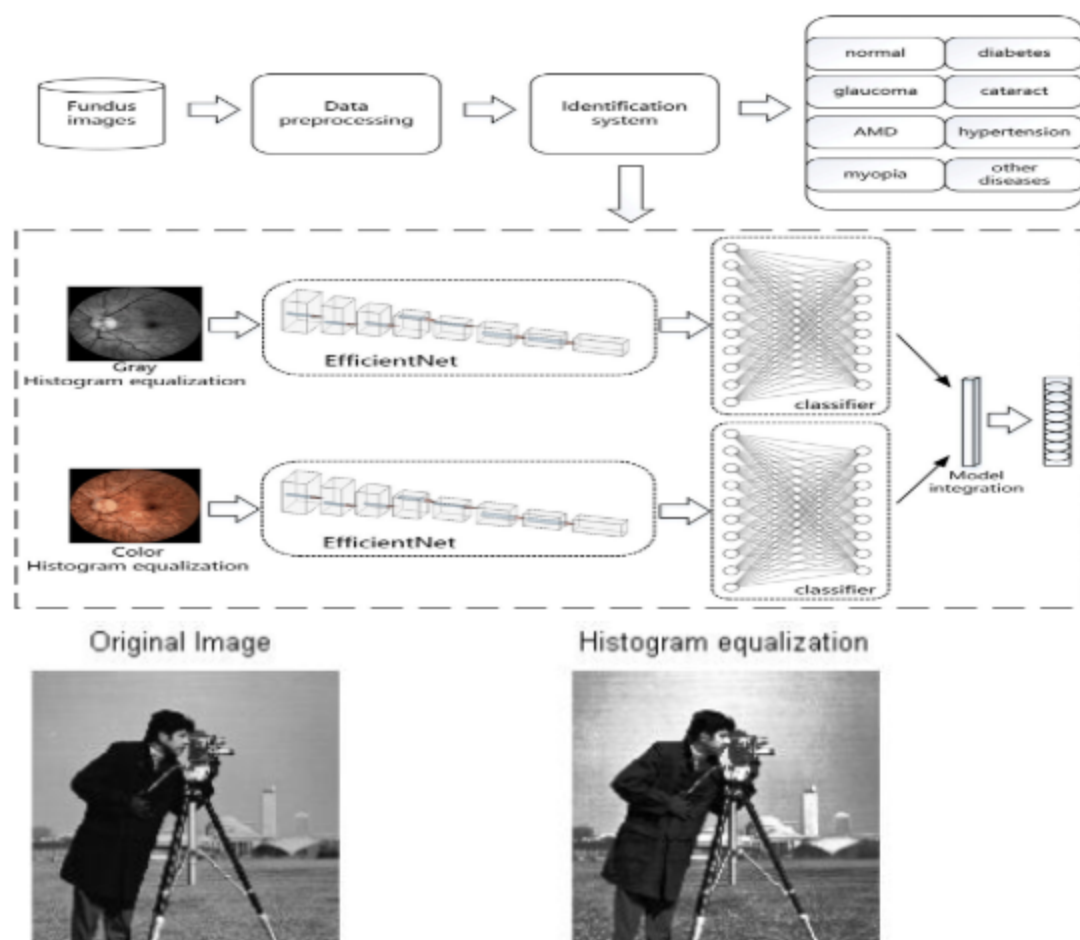
histogram equalization (명암 대비 향상)



Multi-Label Classification of Fundus Images With EfficientNet



논문 리뷰(2)



앙상블(ensemble)

강한 분류 모델 => 2개의 약한 분류모델 결합

원본 + gray 이미지에 histogram equalization 진행

EfficientNet 모델을 통해 학습 (독립)

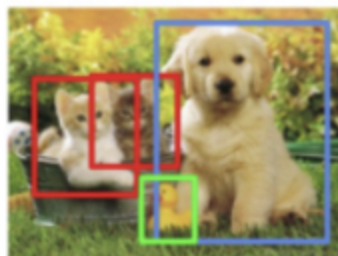
sigmoid 출력 확률을 평균화 -> 최종 출력 값

Histogram equalization?

이미지의 명암을 대비시키기 위해 사용
0-255 사이의 값으로 pixel 값을 재조정
pixel 값 균등 -> 세부사항 더 잘 관찰

논문 리뷰(2)

Object Detection



Instance Segmentation



Configuration	Value
Optimisation function	Adam
Epoch	30(Complete training) 10(fine tuning)
BatchSize	15
Learning rate	1.00E-03
Batch Normalization	True
Drop out	5.00E-01
ReduceLROnPlateau	Monitor='final_score', factor=0.2, Patience=4 ,min_delta=0.001
EarlyStopping	Monitor = 'val_loss', patience=5
ModelCheckpoint	Monitor = 'final_score', mode='Max', save best only=True

다중 레이블 → 일반적인 손실함수 사용 X

다중 손실함수 : 이진 교차 엔트로피 손실함수(Binary Cross-Entropy Loss)

$$L = -\frac{1}{N} \sum_{i=1}^N y_i \log(p(y_i)) + (1 - y_i) \log(1 - p(y_i))$$

N: 샘플 수

y_i : 샘플 i 의 레이블 값

양성 1, 음성 0

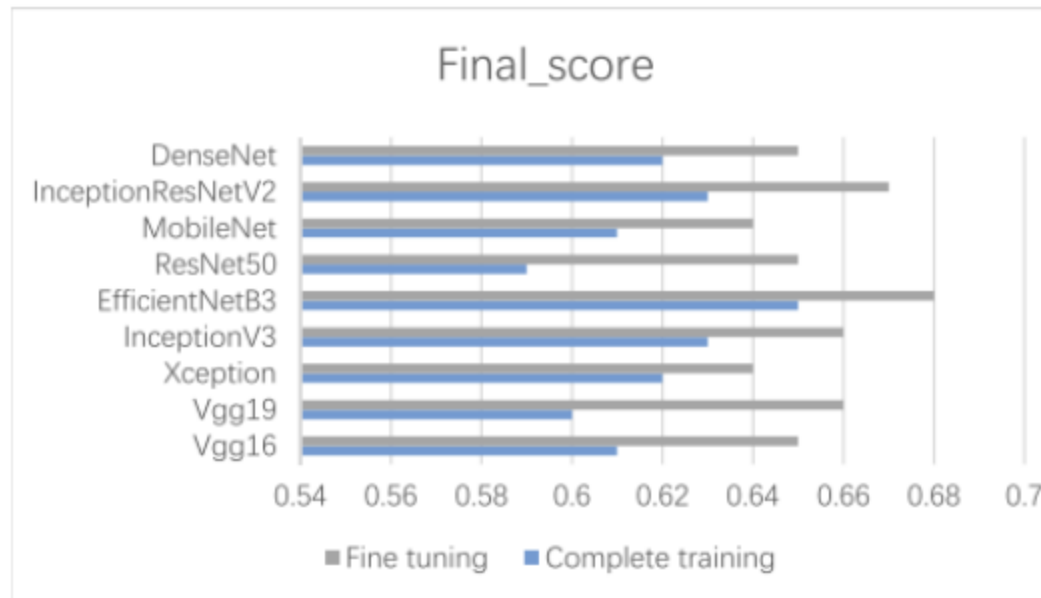
$p(y_i)$: 양성의 값이 나올 확률

SGD // Adam, RMSprop (과거의 기울기 값을 활용)

⇒ 경사하강법 최적화 알고리즘

$$\text{Final_score} = \frac{F1_score + Kappa + Auc}{3}$$

논문 리뷰(2)



Ensemble				
Model	Auc	Kappa	F1_score	Final_score
Vgg16	0.69	0.45	0.85	0.66
Vgg19	0.69	0.45	0.86	0.66
Xception	0.68	0.43	0.85	0.65
InceptionV3	0.68	0.47	0.87	0.67
EfficientNetB3	0.73	0.50	0.88	0.70
ResNet50	0.66	0.42	0.85	0.64
MobileNet	0.67	0.39	0.87	0.64
InceptionResNetV2	0.72	0.47	0.87	0.69
DenseNet	0.70	0.44	0.85	0.66

448x448								
Model	Val_Accuracy	Val_Precision	Val_Recall	Val_Fβ_score	Auc	Kappa	F1_score	Final_score
Vgg16	0.91	0.68	0.58	0.91	0.72	0.50	0.89	0.70
Vgg19	0.91	0.70	0.57	0.91	0.70	0.48	0.88	0.69
Xception	0.92	0.70	0.64	0.92	0.73	0.51	0.89	0.71
InceptionV3	0.91	0.68	0.64	0.91	0.72	0.46	0.87	0.68
EfficientNetB3	0.92	0.71	0.66	0.92	0.74	0.52	0.89	0.72
ResNet50	0.89	0.65	0.56	0.90	0.67	0.45	0.84	0.65
InceptionResNetV2	0.91	0.71	0.64	0.91	0.72	0.49	0.88	0.70
DenseNet	0.91	0.69	0.60	0.91	0.70	0.45	0.87	0.67

final_score, Ensemble, 448x448 등의
여러 실험을 거쳤을 때 가장 좋은 성능을 보임

한계: Network는 이미지 자동으로 특징 학습
→ 구체적인 특징 알 수 없음
이미지 말고도 가족 질환 등의 다양한 요소들이 많음

앞으로 방향성

EfficientNet이 외에 사용할 수 있는 모델 조사

안저 질환에 영향을 미칠 수 있는 요인들에 대한
추가적인 분석

간이 진단에 대한 약처방과 관련된
의약품 데이터 연결 방안 모색



감사합니다.