
ROBUST SKIN DISEASE CLASSIFICATION BY DISTILLING DEEP NEURAL NETWORK ENSEMBLE FOR THE MOBILE DIAGNOSIS OF HERPES ZOSTER

"대상포진 모바일 진단을 위한 딥러닝 앙상블 지식을 활용한 피부 질환 분류 기법"

3조 [REDACTED] 최우연

목표 연구과제 : "CNN기반 대상포진 조기진단 AI 모델 구축"

1. 논문 출처 및 핵심 내용

논문 출처



- IEEE(세계 최대 전기전자공학 학회)가 발행하는 오픈 액세스 (Open Access) 학술지에 게재

- 논문 게재일 : 2021년 1월 25일
- 논문 저자 : 본 연구는 광주과학기술원(GIST), 유니트리아, 연세대학교 의료원 소속 연구진 10명이 공동 수행

핵심 내용

- 모바일로 촬영한 이미지에서도 대상포진을 정확하게 진단하기 위해 여러 딥러닝 모델을 압축해 만든 가볍고 튼튼한 AI 모델을 개발한 것
- 문제 : 모바일 사진은 흔들림, 노이즈 등으로 품질이 낮아 진단이 어려움
- 목표 : 모바일에서도 잘 작동하는 가벼운 딥러닝 모델 개발
- 방법 : 강한 앙상블 모델에서 학생 모델로 지식을 점진적으로 전이(KDE-CT)시키는 작고 가벼운 모델 개발
- 성과 : 높은 정확도(93.5%)와 낮은 연산량으로 모바일 진단에 최적화된 결과 도출

2. 연구 배경

대상 포진의 위험성

- 통증과 수포를 동반하는 바이러스성 피부 질환
- 평생 발병 확률 10%~30%
- 72시간 내 항바이러스 치료 미 제공 시 만성 질환으로 진행 가능
- 삶의 질 저하, 수면, 사회 활동에 부정적 영향

자동 진단 시스템의 필요성

- 피부질환 대부분의 연구는 흑색종 중심
- 기존 연구는 피부확대경 이미지만 활용
- 모바일 기반 임상 이미지 진단 시스템 부재
- 사용자가 촬영한 임상 이미지를 통한 자가 진단
- 서버 업로드 없이 기기 내 오프라인 진단 필요
- 다양한 조명과 노이즈에 강한 견고한 모델 필요

3. 연구 목표 및 차별점

기존 연구의 한계

- 대규모 모델(Inception-V4, 앙상블 네트워크) 사용 -> 모바일 적용의 어려움
- 모델 최적화를 위한 신경망 구조 탐색(NAS) 방법의 복잡성 존재
- 시각적 변형에 대한 강건성 평가 시 노이즈나 블러에만 초점을 맞춘 제한적인 평가가 이루어짐

본 연구의 차별점

- 임상 이미지를 활용하여 대상포진 진단용 모바일 모델 개발
- 지식 증류(KD) 기반 경량화 및 성능 최적화
- 75종의 시각적 변형에 대한 견고성 평가
- MobileNetV3-Small 기반으로 93.5% 정확도, 67.6 평균 손상 오류 달성

4. 연구 차별점 - 시각적 변형에 대한 모델 견고성

KDE-CT 모델 훈련기법 적용에 의한 강건성 비교

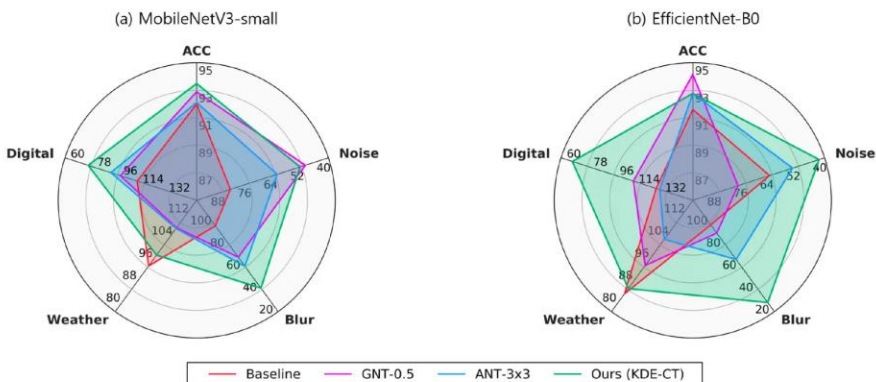


FIGURE 1. Summary of key results. Our knowledge distillation from ensemble via curriculum training (KDE-CT) enhanced the corruption robustness for both (a) MobileNetV3-Small and (b) EfficientNet-B0 compared to the standard training (baseline), Gaussian noise training (GNT), and adversarial noise training (ANT). ACC is the overall accuracy, and the rest are the corruption errors for noise, blur, weather, and digital corruptions.

이미지 분류 딥러닝 모델 비교

- MobileNetV3-small : 모바일.임베디드 기기용의 경량모델, 빠르고 가볍음
- EfficientNet-B0 : 효율성과 정확도의 균형을 맞춘 모델로 높은 정확도를 보이나 상대적으로 무거움

모델 훈련방법 비교

- Baseline : 데이터 증강이나 증류 기법 없이 일반적인 방식으로 훈련된 모델
- GNT, ANT : 각각 랜덤 노이즈 또는 적대적 노이즈를 통해 모델을 견고하게 만드는 기법
- KDE-CT : 여러 모델의 지식을 활용해 단계적으로 학습시키는 기법

결론

- 모바일 환경에 적용 가능한 경량 모델인 MobileNetV3-small에 KDE-CT 훈련 기법을 적용함으로써, 모델의 정확도와 시각적 변형에 대한 견고성을 크게 향상시킴

5. 연구 방법론 - 데이터 세트

SD-HZ : 대상포진 진단용 피부 질환 이미지 데이터 세트 구성



FIGURE 2. Examples images in SD-HZ dataset: (a) Acne (b) Herpes Zoster (c) Tinea (d) Other Disease.

공개 데이터 세트

- SD-198(6,548장, 198개 질환 클래스) + SD-260(20,600장, 260개 질환 클래스) 수집
 - 모바일 기기 및 디지털 카메라로 다양한 조명환경과 카메라 설정 하에서 다양한 피부 타입을 대상으로 촬영한 이미지 포함
 - 단, 두 데이터 세트의 대상 포진 이미지는 각각 24장, 12장으로 학습에 부족함



커스텀 데이터 세트

- Google/Bing에서 'herpes zoster' 키워드로 웹 크롤링
- 중복 제거 및 전문가 검토를 통해 총 377장의 대상포진 이미지 확보

5. 연구 방법론 - 데이터 세트

SD-HZ-C : 손상 시뮬레이션 이미지 데이터 세트 구성



FIGURE 3. Examples of HZ image in HZ-W where light condition and camera configurations are varied (from close-up to long shot).

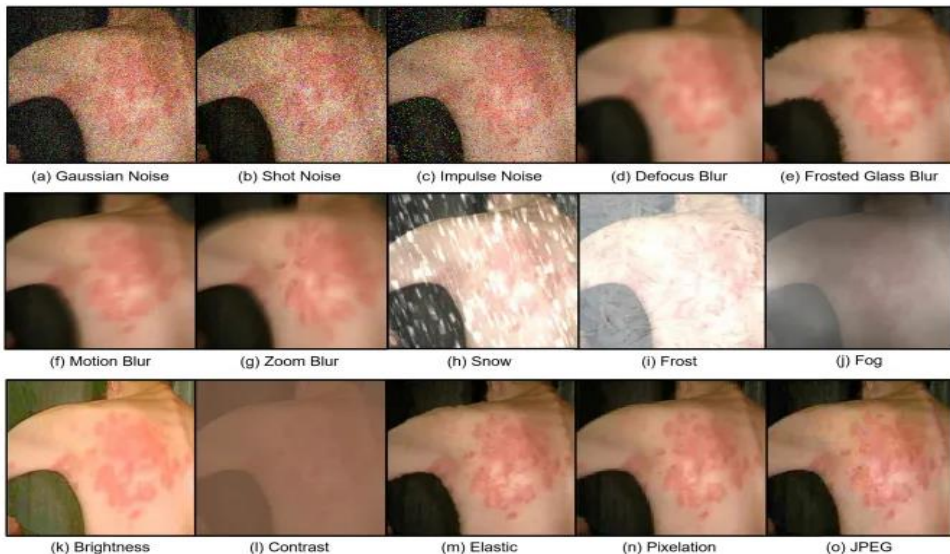


FIGURE 4. Examples of corrupted images in SD-HZ-C owing to (a)-(c) noise, (d)-(g) blur, (h)-(j) weather, and (k)-(o) digital categories with severity level of 3.

강건성 평가를 위한 데이터 세트 생성

- 사용자 업로드 이미지의 손상 상황을 고려하여 모델의 강건성을 평가하기 위해 데이터 세트를 생성
- SD-HZ 데이터 세트의 테스트 세트에 총 75가지 손상 적용
- 15개 손상유형 X 5단계 심각도 레벨로 구성
 - 노이즈 : 가우시안 노이즈, 샷 노이즈, 임펄스 노이즈
 - 블러 : 초점 흐림, 얼음 유리 흐림, 모션 블러, 줌 블러
 - 날씨 : 눈, 서리, 안개
 - 디지털 손상 : 밝기 변화, 대비, 탄성 왜곡, 픽셀화, JPEG 압축

6. 연구 방법론 - 모델 학습 전략 및 알고리즘 설계

KDE-CT 학습 및 테스트 과정

Training 과정

- Teacher Networks
 - 크고 정확한 모델들로 학습
 - 깨끗한 이미지로 학습
 - Knowledge Distillation : Student network에 지식전달
- Student Network
 - 모바일에서 사용할 수 있는 작고 가벼운 모델
 - 여러 Teacher Network로 부터 학습함
- KDE-CT
 - 점진적으로 학습강도를 높이는 방식
 - 학습 초반에는 쉬운 이미지, 이후에는 어려운 이미지로 학습

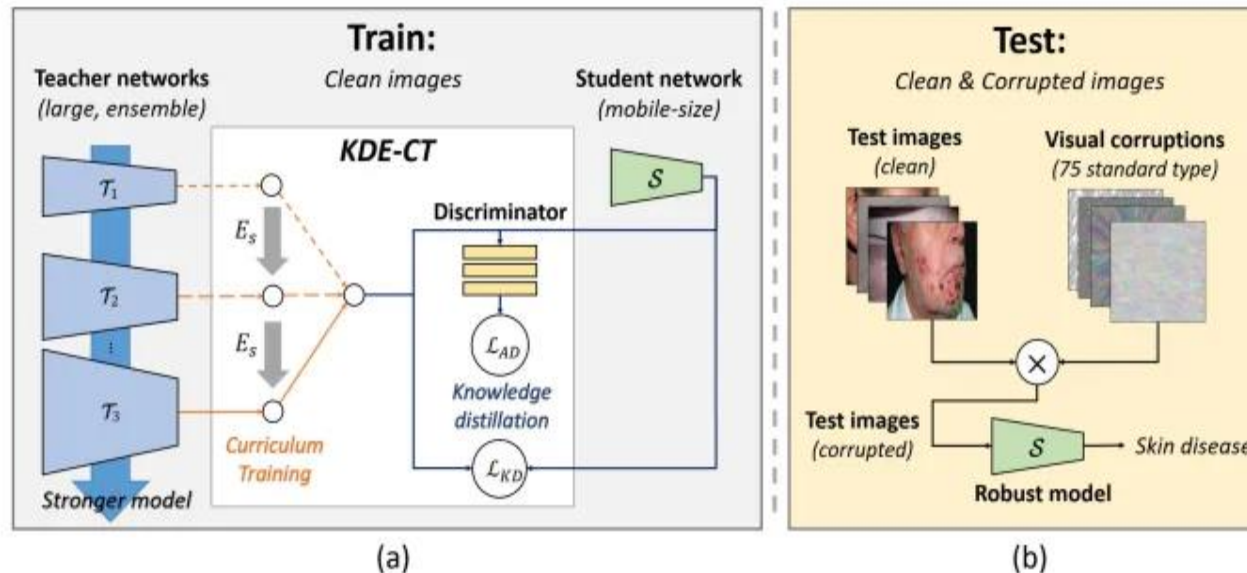


FIGURE 5. Flowchart of proposed KDE-CT. For the diagnosis of HZ from a clinical image, (a) we train a mobile neural network (S) on clean images, and then (b) test this model on both clean and corrupted images, where the corrupted images are generated with 75 standard visual corruptions (e.g., Gaussian noise, zoom blur, fog, and pixelation). The student network (S) learns from an ensemble of large teacher networks (T_1, T_2, \dots, T_N) using KD from ensemble via curriculum training (KDE-CT) by minimizing the sum of distillation loss and adversarial loss. The teacher network for training is progressively updated for every E_s epoch so that the student network can learn more robust and richer features from multiple teacher networks under a better curriculum.

Test 과정

- 학습된 모델인 Student Network를 실제 이미지에 적용하여 진단
- 테스트 이미지는 두 종류
 - 깨끗한 이미지
 - 손상된 이미지 : 노이즈, 블러, 픽셀화 등 총 75가지 시각적 손상
- Student model은 이 두 유형의 이미지를 모두 입력받아 피부 질환 진단 수행
- 결과적으로 깨끗하고 손상된 이미지 모두에서 잘 작동하는 강건한 Model 획득

7. 연구 방법론 - 지식 증류 방법론

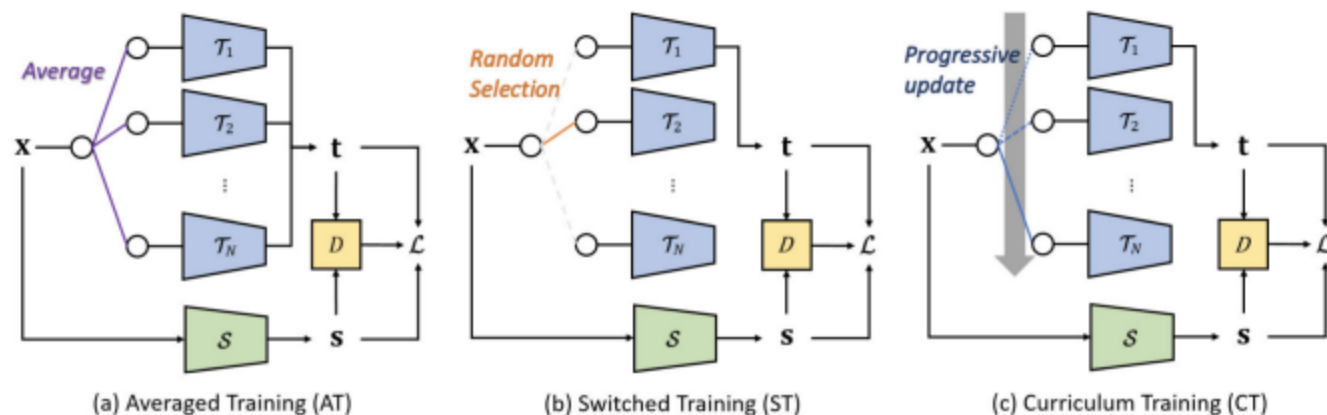


FIGURE 6. Illustration of training strategies for the knowledge distillation from DNN ensemble. S is student network and T is teacher network. To distill the knowledge of multiple teacher networks, (a) the output of all teacher models is averaged in AT, (b) a random teacher for each batch is selected in ST, and (c) a teacher is progressively updated for every (e_c) epoch in CT.

KDE-AT	KDE-ST	KDE-CT
<ul style="list-style-type: none"> • KDE via Averaged Training • 가장 단순한 형태의 앙상블 지식증류 방식 • 여러 Teacher 모델들의 예측결과를 평균내어 Student 모델에게 전달 • Teacher 모델들의 다양성을 살리지 못해 이미지 왜곡에 취약함 	<ul style="list-style-type: none"> • KDE via Switched Training • Training 과정에서 각 미니배치마다 무작위로 하나의 Teacher 모델을 선택 • 해당 Teacher의 예측 결과를 기반으로 Student 모델을 학습 • 학습의 일관성과 Teacher 통합 효과가 부족해 견고성 향상에 한계가 있음 	<ul style="list-style-type: none"> • KDE via Curriculum Training • 성능이 낮은 Teacher 모델부터 높은 Teacher 모델 순서로 점진적으로 Student 모델을 학습시키는 방식 • Student 모델을 점차 정교하게 학습시키는 커리큘럼 기반 증류방법 • 학습 안정성과 견고성을 향상시킴

8. 실험 설정 및 평가 지표

데이터 구성

➤ 데이터 분할

구분	비율	설명
Train	70%	ImageNet 사전학습->FC 레이어 수정
Val	10%	Epoch별 검증 손실 모니터링
Test	20%	SD-HZ, SD-HZ-C 테스트 세트 평가

- Stratified Sampling 사용 (클래스 비율 유지)

➤ 데이터 증강

- 라이브러리 : Albumentations
- 이미지 크기 : 224 x 224
- 적용 기법
 - 좌우/상하 반전
 - 밝기/대비/색도/채도 조정
 - 이동, 회전, 스케일 변환

모델 훈련 및 평가

➤ 모델 비교 평가

- 총 13개 모델 평가
- 기본, 모바일, 앙상블 모델 평가
- 모델 일반화 성능 평가시 데이터셋은 SD-HZ(깨끗한 이미지 기준)으로 평가
- 모델 강건성 평가시 데이터셋은 SD-HZ-C(손상된 이미지 기준)으로 평가

➤ 학습 세부 설정

항목	기본 모델	KDE 기반 모델
Optimizer	RAAdam	SGD
Batch Size	64	128
Learning Rate	2×10^{-4}	$\mathcal{L}_{KD} = 0.1$, $\mathcal{L}_D = 0.001$
Epochs	Early Stopping (10 epoch 기준)	최대 500 epoch, 150/250 epoch 감쇠 적용

평가 지표

➤ 성능 지표

ACC	전체 정확도
F1, PR, RE	F1-score, 정밀도, 재현율의 Macro평균
AUROC	ROC 커브 하단 면적(모델 판별 능력)
F1-HZ	대상포진(HZ) 클래스의 F1 점수
ACC-HZ	대상포진(HZ) 클래스의 정확도
Kappa	전문가와 모델 간 예측 일치도

➤ 계산 효율성 지표

MACs	Multiply-Accumulate 연산 수(계산량)
Latency	이미지 1장을 처리하는 시간(테스트 지연시간)
Params	파라미터 개수(모델 저장 크기)

➤ 강건성 평가 지표

CE	손상 유형별 오류율
mCE	15개 손상 유형에 대한 CE의 평균값

9. 실험 결과

1. 깨끗한 이미지(SD-HZ)에서의 분류 성능

TABLE 2. Performances of 13 different models on the SD-HZ dataset (Latency: millisecond, Params.: million, MACs: million).

Category	Model	ACC	F1	PR	RE	F1-HZ	AUROC	Kappa	Latency	Params.	MACs
Basic models	AlexNet	0.93	0.84	0.87	0.82	0.80	0.946	0.794	21.0	57.0	710
	Vgg-16	0.93	0.85	0.87	0.86	0.79	0.967	0.779	140.3	134.3	1,550
	InceptionV3	0.92	0.83	0.85	0.82	0.78	0.960	0.766	73.2	21.8	285
	ResNet-50	0.94	0.87	0.87	0.88	0.81	0.979	0.797	93.3	23.5	412
	ResNext-101	0.93	0.87	0.87	0.89	0.82	0.975	0.809	320.7	86.8	1,651
	SEResNext-101	0.94	0.88	0.89	0.87	0.85	0.981	0.846	226.8	46.9	805
Mobile models	DenseNet-121	0.93	0.85	0.87	0.84	0.80	0.969	0.791	328.6	7.0	285
	EfficientNet-b0	0.92	0.81	0.82	0.83	0.68	0.954	0.658	55.5	4.0	40
	MNasNet	0.92	0.84	0.85	0.84	0.76	0.947	0.749	38.6	3.1	33
	MobileNetV2	0.90	0.79	0.80	0.83	0.61	0.955	0.586	39.5	2.2	32
	MobileNetV3-Small	0.92	0.84	0.85	0.83	0.82	0.927	0.811	20.6	1.5	5
DNN Ensemble	Ensemble-M	0.95	0.89	0.92	0.88	0.82	0.979	0.880	666.7	142.2	2,746
	Ensemble-E	0.97	0.93	0.94	0.92	0.78	0.992	0.910	699.5	144.6	2,781

- DNN 앙상블 모델인 Ensemble-E는 모든 모델 중 가장 높은 정확도 (ACC 0.97), F1-score(0.93), AUROC(0.992)를 기록하며, 정상이미지에 대한 분류 성능이 가장 우수함
 - 다만, 연산량과 모델 크기가 크고, 모바일 환경에 적합하지 않음
- MobileNetV3-small은 경량모델임에도 불구하고 ACC 0.92, F1-HZ 0.82로 양호한 성능을 보임
 - MobileNetV3-small은 파라미터수와 연산량, 지연시간이 매우 낮아 모바일 환경에 적합한 경량 구조임

2. 손상된 이미지(SD-HZ-C)에서의 견고성 평가

TABLE 4. Robustness of 13 different models on corrupted images (SD-HZ-C), where the models are trained using only clean images (SD-HZ). Here, ACC is the overall accuracy (higher is better). The values in Noise, Blur, Weather, and Digital columns are corruption errors, and mCE is the mean of these values (lower is better).

Model	ACC		Noise			Blur				Weather				Digital			
	ACC	mCE	Gauss.	Shot	Impulse	Defocus	Glass	Motion	Zoom	Snow	Frost	Fog	Bright	Contrast	Elastic	Pixel	JPEG
AlexNet	92.8	100.0	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
Vgg-16	93.2	124.1	44	116	169	116	148	125	149	89	130	116	94	101	152	174	140
InceptionV3	91.8	79.2	31	77	74	42	63	105	49	108	96	57	93	47	124	125	96
ResNet-50	93.7	106.6	50	142	128	97	117	124	140	85	123	92	90	46	185	91	90
ResNext-101	93.4	132.0	99	248	251	111	107	187	143	93	125	149	70	71	119	106	101
SEResNext-101	94.1	87.2	34	96	80	51	76	100	106	86	94	91	83	73	119	130	88
DenseNet-121	93.4	116.6	49	143	148	116	121	107	129	84	85	115	87	89	170	161	146
EfficientNet-B0	91.6	96.2	33	83	79	75	86	104	127	82	98	79	97	44	149	190	117
MNasNet-B1	92.4	100.3	59	132	129	35	69	78	92	104	133	70	124	60	146	158	115
MobileNetV2	90.0	142.2	74	191	184	97	144	161	91	105	113	84	149	80	225	273	163
MobileNetV3-Small	92.0	98.6	43	102	109	67	62	149	110	107	115	67	110	78	125	125	110
Ensemble-M	95.0	81.1	34	89	85	70	79	90	115	87	73	84	64	77	101	92	75
Ensemble-E	96.5	69.7	27	83	63	66	74	65	96	70	62	87	48	65	91	84	65

- DNN 앙상블 모델인 Ensemble-E는 손상된 이미지에서도 ACC 96.5%, mCE 69.7로, 시각적 변형(노이즈, 블러, 픽셀화 등)에 대해 가장 높은 견고성을 보임
- 반면, MobileNetV3-small은 ACC 90.2%, mCE 98.6으로 견고성 면에서는 시각적 손상에 취약한 결과를 나타냄
 - 단순 모델 구조만으로는 시각적 손상 대응에 한계가 있음을 발견
 - 따라서, MobileNet기반 모델에 KDE-CT와 같은 지식 증류 기반 견고성 향상 기법을 적용하는 것이 필수적임을 확인

10. 실험 결과 및 최종 결론

3. 지식증류 + 커리큘럼 학습기법에 의한 견고성 향상

TABLE 6. Comparison of different methods to improve the robustness against common visual corruptions (SD-HZ-C), where the models are trained using only clean images (SD-HZ). Here, ACC is the overall accuracy (higher is better). The values in the Noise, Blur, Weather, and Digital columns are corruption errors, and the mCE is the mean of these values (lower is better). The underlined KDE-CT significantly enhanced the corruption robustness of the models when compared with other methods.

		(%)		Noise			Blur				Weather				Digital			
Architecture	Method	ACC	mCE	Gauss.	Shot	Impulse	Defocus	Glass	Motion	Zoom	Snow	Frost	Fog	Bright	Contrast	Elastic	Pixel	JPEG
MobileNetV3-Small	Baseline	92.0	98.6	43	102	109	67	62	149	110	107	115	67	110	78	125	125	110
	GNT $\sigma_{0.5}$	92.9	83.7	25	63	63	53	36	104	81	113	111	107	106	57	127	104	103
	ANT $\sigma_{0.3}$	92.1	81.7	31	75	84	28	41	93	80	102	113	115	98	59	108	101	98
	KD-SM	93.2	85.4	25	64	63	68	50	99	102	94	121	72	102	104	87	122	108
	KDE-AT	93.4	78.6	31	92	77	52	37	78	90	96	95	109	94	118	79	78	52
	KDE-ST	93.2	68.0	28	77	73	38	24	58	77	92	91	81	77	114	71	74	44
	KDE-CT	93.5	67.6	25	68	66	28	28	61	48	101	97	102	86	109	76	75	44
EfficientNet-B0	Baseline	91.6	96.2	33	83	79	75	86	104	127	82	98	79	97	44	149	190	117
	GNT $\sigma_{0.5}$	94.1	94.5	41	98	100	80	81	86	115	88	77	125	91	76	124	149	88
	ANT $\sigma_{0.3}$	92.8	90.4	26	77	60	44	61	81	81	85	113	119	98	76	138	171	124
	KD-SM	93.3	70.1	33	90	89	25	25	62	44	97	95	81	110	106	76	76	42
	KDE-AT	92.4	85.6	49	143	126	31	32	66	67	95	104	83	101	125	90	102	69
	KDE-ST	92.8	67.2	24	63	58	26	27	64	49	90	102	106	85	106	79	83	45
	KDE-CT	92.8	56.7	19	56	52	17	18	51	27	78	98	88	74	80	68	78	44

- KDE-CT 방식이 두 모델(MobileNetV3-small, EfficientNet-B0) 모두에서 가장 낮은 mCE를 기록
 - 손상된 이미지(노이즈, 블러, 픽셀 등)에 대한 견고성 향상 효과가 가장 큼
- KDE-CT 방식 적용시 ACC도 상승됨
 - 단순히 강건할 뿐만 아니라 정확도도 개선됨
- KDE-CT는 경량 모델과 고성능 모델 모두에 적용 가능하고, 손상 이미지 상황에서도 가장 높은 견고성을 확보하는 데 효과적인 학습전략임을 입증함

Conclusion ...

- ✓ 모바일 피부 질환 모델 선택 시, 정확도 뿐만 아니라 시각적 손상에 대한 강건성도 함께 고려해야 함
- ✓ 경량화된 모델에 지식 증류 훈련방법을 적용하여 강건하고 뛰어난 성능 개선이 가능함
- ✓ KDE-CT는 모바일 피부 진단에 적합한 정확하고 강건한 경량 DNN을 학습할 수 있는 효과적인 방법임을 연구에서 입증함

Thank you
