**4. Experiments**

**-실험 환경 및 Datasets**

* Pick-a-back 모델: 4개의 유명 아키텍처에서 4개의 다른 분류 작업 그룹과 13개의 공개 데이터셋을 사용하여 검증
* 사용된 데이터셋: CIFAR-100, Classified-ImageNet, Fine-Grained, Digits. *부록 섹션 A1*

**-Architectures**

* 경량 모델: LeNet-5, MobileNetV1 (ImageNet 사전 학습된 모델 없음)
* 대규모 모델: VGG16-BN, EfficientNet-B0 (ImageNet 사전 학습된 모델 있음, PyTorch로 구현)
* 기본적으로 CIFAR-100을 LeNet-5에서 사용

텍스트, 스크린샷, 마조렐 블루, 라인이(가) 표시된 사진

자동 생성된 설명

(a)Absolute Acc. (a)절대적인 정확도

(b)Improved Acc. 향상된 정확도

**Fig. 2 (a, b) 설명**:

* **(a) 절대 정확도**: 백본(Backbone)과 목표 작업(Target Task)에 대한 절대 정확도

X축: 백본 작업 / Y축: 목표 작업

첫 번째 행은 모든 작업에 대해 평균 값을 나타냄

첫 번째 열은 백본이 없는 경우의 결과를 보여줌

* **(b) 정확도 향상**: 백본 없이 진행된 학습과 비교했을 때 … 각 백본에 대한 목표 작업의 정확도 향상 정도

절대 값이 아닌 정확도의 개선을 기준으로 시각화됨

**Fig. 3 (a, b) 설명**:

* **(a) 학습 동적 과정**: 4번 목표 작업에서 백본 작업을 활용한 학습 동적 과정을 나타냄

여기서 백본 없이 진행한 학습과 Pick-a-back 기법을 사용한 경우의 훈련 및 테스트 정확도가 시간에 따라 어떻게 변하는지 볼 수 있음

* **(b) 정확도 개선**: 가장 성능이 좋은 목표 작업인 14번 작업에 대한 정확도 개선을 보여줌

백본 없이 진행한 학습과 비교하여 Pick-a-back을 통해 개선된 정확도를 스무딩 처리한 결과

X축의 두 번째 값은 가장 낮은 성능을 보인 두 가지 작업의 평균 값을 나타냄

마지막 값은 모든 작업의 평균

**-Baselines(비교 대상)**

* Pick-a-back은 다양한 연합 학습, 지속 학습, 지속 연합 학습 접근법과 비교됨
* **연속 학습 기법과 비교**

ProgressiveNet(순차적 작업에 대한 스크래치 모델을 연결하여 확장해야만 하는 연속 학습)

PackNet(각 작업에만 부분 뉴런을 할당하여 가지치기만 하는 연속 학습)

CPG(필요한 경우 확장 및 모델 소형화를 위한 가지치기가 포함된 연속 학습)

WSN(이전 하위 네트워크의 마스킹된 가중치를 재사용하여 지속적으로 학습)

* **연합 학습 기법과 비교**

FedAVG(연합 학습에서 모델 정규화를 위한 선구적인 작업 중 1)

FedWeIT(글로벌 연합 매개변수와 함께 희소한 작업별 매개변수를 도입하여 연합 지속 학습 패러다임)

FCCL(이질적인 환경에서 치명적인 망각 문제를 해결하기 위한 연합 지속 학습 접근 방식)

* **지식 선택 전략과 비교**

백본: 선택적 지식 전이 X, 손상 X 연속 학습

균일한 지식: 백본 모델에 동일한 작업을 적용+각 작업에 대한 모든 결과를 평균화함

내부 지식: 대상 작업을 백본 모델로 반복적으로 훈련

무작위 지식 선택: 백본 모델을 무작위 선택

이상적인 상위 모델: 가장 높은 정확도 향상을 보이는 특정 작업을 가진 이상적인 백본을 선택

**-Measures(측정 지표)**

* **Avg**: 모든 작업의 평균 정확도 (**정확도**)
* **Top**: 선택적 지식 전이를 통해얻은 가장 높은 정확도 (**정확도**)
* **모델 크기**: 모델의 실제 저장 용량-오버헤드(**타당성**)
* **확장성**: 스크래치 모델의 값이 1인 확장 정도 … 모델이 얼마나 확장되었는지 측정
* **Sparsity (희소성)**: 모델의 희소성 비율 … 희소성이 높을수록 다음 작업을 위한 공간이 더 많음을 의미
* **Training time (훈련 시간)**: 한 에포크(시기)를 실행하는 데 소요되는 시간
* **Used data (사용된 데이터)**: 훈련에 사용된 데이터 샘플 수

표 1: 이 표는 **성능 비교**를 나타냄

각 클라이언트(15, 1, 17 등)의 모델 성능을 비교함. 표의 굵은 글씨는 각 클라이언트에서 가장 높은 성능을 보인 항목을 강조하고 있음

1. **Client ID**: 각 클라이언트(디바이스) ex) 15, 1, 17 등으로 구분됨
2. **w/o Backbone**: 선택적 지식 전이 없이 학습한 경우의 성능 (기본 성능)
3. **다양한 모델 학습 기법**: FedAVG, ProgressiveNet, PackNet, WSN, FedWeIT, FCCL
4. **Pick-a-back**: 마지막 행 주목 !!! Pick-a-back 모델의 성능
5. **Upper**: 이론적으로 가능하지만 실제로는 불가능한 최적의 성능 - 상한선으로 사용

* **FedAVG와 FCCL**: 기존의 연합 학습 및 연합 지속 학습 기법

성능이 비교적 고르게 분포되어 있지만, Pick-a-back보다 전반적으로 낮음

* **ProgressiveNet, PackNet**: 지속 학습 기법 중 일부는 특정 클라이언트에서 좋은 성능을 보임(전반적으로는 Pick-a-back보다 낮음)
  + **Avg (평균)**: 마지막 열 맨끝 (**Pick-a-back):** 71.8 ->대부분의 기존 학습 기법보다 good

즉 Pick-a-back은 다양한 기법 중에서 대부분의 클라이언트에서 성능이 우수하고 기존 기법을 사용하는 것보다 **효율적이고 성능이 뛰어남**

**4.1 Overall Performance and Comparisons**

**-선택적 지식 연합의 동기**

백본(backbone) 사용을 통해 연합된 지식: 지속 학습자에게 이점을 줌

하지만 일부 부정적 백본도 존재하여 목표 과업에 맞춘 선택적 지식이 필요함

성능이 낮은 과업은 더 큰 정확도 향상을 보이는 반면, 잘 구성된 모델은 외부 지식에 의해 방해받을 가능성이 있음

**-선택적 지식 연합의 효과성 검증**

다양한 학습 방법과 백본 선택 전략을 비교한 결과, Pick-a-back이 20개 과업 중 2개를 제외하고 평균적으로 뛰어난 성능을 보였음.( 선택 접근 방식을 사용하여 최대 8.0%의 상당한 정확도 향상을 관찰)

FedAVG는 지속 학습 과업을 고려하지X -> 학습한 지식을 잃는 문제

Pick-a-back은 다양한 방법(ProgressiveNet, PackNet)에서도 성능을 개선

* + 백본 선택 전략과 함께 지식 재사용 접근 방식의 효과 때문

**-백본 지식의 이점**

Pick-a-back은 훈련 초기부터 테스트 정확도가 w/o Backbone을 능가

* + 간접 지식이 개별 학습자가 경험이 부족한 지식에 노출되고 보이지 않는 데이터에 적응하는 데 도움이됨

지식 전이를 통해 새로운 데이터를 더 빨리 학습함

Pick-a-back은 20 에포크 내에 w/o Backbone이 50 에포크 동안 얻는 성능을 도달할 수 있음 -> 학습 속도를 가속화하고 전반적인 성능을 향상시킴

**-유사하지만 일반화된 지식의 가능성**

동일한 백본을 모든 과업에 적용한 Uniform 지식과 비교했을 때 Pick-a-back은 초기 성능이 낮은 과업에서 더 큰 개선을 보였음

단일 과업을 반복적으로 학습하는 내부 지식은 성능을 더욱 향상시키지만

Pick-a-back은 유사한 간접 경험을 활용하여 더 나은 성과를 달성함.

* + 개별 과제를 넘어서는 일반화와 더 나은 개선을 제공한다는 것을 의미

**4.2 Scalability Analysis**

**-엣지 AI를 위한 경량 모델**

LeNet-5, MobileNet-V1, EfficientNet-B0와 같은 **경량 모델에서** Pick-a-back을 평가

* + 최대 9.8% 정확도 향상

대규모 모델(VGG16-BN): 엣지 환경에서 메모리와 처리 능력 제한->배포가 어렵

경량 모델: 서버 의존성을 줄임, 빠른 학습과 추론 시간을 제공

Pick-a-back: 경량 모델에서도 큰 성능 개선을 이룸

저장 공간과 실행 시간을 크게 줄일 수 있음

표 2: **CIFAR-100 데이터셋에서 다양한 아키텍처에 대한 정확도, 모델 크기, 평균 학습 시간**

Pick-a-back을 사용하면 모든 모델에서 정확도가 향상됨

특히 VGG16-BN과 EfficientNet-B0에서 성능 차이가 큼

모델 크기와 학습 시간도 각 모델에 따라 차이가 있지만 모두 Pick-a-back 사용에 따른 성능 향상을 보임!!

표 3: **LeNet-5 모델을 사용한 다양한 데이터셋 성능 비교**

**Pick-a-back 사용 시 모든 데이터셋에서 성능이 향상됨** 특히 CIFAR-100과 Classified-ImageNet에서는 최고 성능에서 각각 8.0%와 8.4%

**Fine-Grained 데이터셋**에서는 평균 성능과 최고 성능 모두 개선되었지만, 다른 데이터셋에 비해 성능 향상 폭이 더 작음

Upper 성능 vs Pick-a-back… 픽어백은 백본 없는 경우보다 항상 우수하지만Upper 성능보다는 약간 낮음

**-다양한 데이터셋에서의 확장성**

Pick-a-back은 다양한 데이터셋에서 안정적으로 성능 개선을 이룸

Digits 데이터셋에서 로마 숫자 과업과 같은 복잡한 과업에서 5.72% 성능 향상

classified-ImageNet에서도 성능이 낮았던 과업에서 큰 개선을 이룸

Fine-Grained 데이터셋에서는 많은 클래스 수로 인해 분류가 어려워짐

But Pick-a-back으로도 성능 개선이 가능

특히, (정교한 분류가 필요한)복잡한 과업에서 Pick-a-back의 개선 잠재력이 큼

**표 4: EfficientNet-B0를 사용한 Fine-Grained 데이터셋 성능**

Pretrained 모델보다 Pick-a-back을 사용할 때 성능이 크게 향상됨

특히 백본 없을 때의 성능(평균에 주목)보다 약 2배 가까이 개선됨..

……………….표 5도 똑같이 픽어백 사용시 성능이 향상됨을 보여줌!(특히 Heterogeneous)

**-가장 복잡한 과업에서의 영향**

CUB, Stanford Cars, Flower, WikiArt, Sketch 같은 복잡한 과업에서 Pick-a-back 적용 후 평균 31.8%의 성능 향상, 최대 51.9%까지 향상됨

ImageNet-사전학습 모델의 한계를 넘어서는 성능을 보여줌

-> 많은 데이터가 필요 … Pick-a-back: 선택적 지식 전이로 이러한 한계를 극복

**-이질적인 데이터 시나리오에서의 우수성**

이질적인 데이터 환경에서 Pick-a-back이 효과적

데이터 이질성이 커질수록 scratch 모델의 성능이 저하

Pick-a-back <- 약한 클래스에서는 외부 지식이 유익

이질적 데이터에서 최대 18.4%까지 성능 향상이 이루어짐

표 6: **VGG16-BN 모델에서 Pick-a-back 성능 비교**

Pretrained 모델 사용 여부에 관계없이 Pick-a-back을 적용하면 성능이 향상됨

특히 Pretrained 없이도 성능 개선 good

학습 데이터 크기도 Pretrained 미사용 시 훨씬 적음 (0.3 GB vs 92.8 GB)

**4.3 Discussions and Limitations**

**-ImageNet 사전학습 모델**

LeNet-5 같은 경량 모델: Pick-a-back을 사용해도 대규모 모델을 능가하지는 못함

VGG16-BN과 같은 ImageNet 사전학습 가중치를 사용한 모델: 시각적 과업에서 강력한 성능 가짐

* + but, 대규모 라벨링된 데이터셋이 필요하고 중앙 서버로 업로드 필요

사전학습 모델: 이미지 분류에 특화됨 -> 특정 또는 전문 과업에 적용할 때 성능 차이가 있을 수 있음

사전학습 모델: 대규모 아키텍처에서만 사용 가능 -> 엣지 학습에 적용하기에는 어려움

사전학습 모델 없이도 Pick-a-back을 사용한 선택적 지식 연합: 평균적으로 원래 사전학습 모델 대비 3.9%의 성능 차이만 보이며, 학습 데이터는 단 0.3%만 사용

Pick-a-back의 지속 학습자와 여러 지속 학습 접근 방식에서의 망각 방지의 직교성을 추가로 제시

이웃 클라이언트 간 이미지 및 모델 공유로 인해 전송 오버헤드 및 개인정보 문제 발생 가능성이 있음 -> 통신 및 계산 시간과 공격 방어 방법을 추가로 평가함