

문제점 개요서

Project Name	Multi-Task Learning의 최적 공유 전략 연구 및 실험적 유효성 분석
-----------------	---

8조

202001156 김수영

202002510 송재현

지도교수: 이종률 교수님 (서명)

Document Revision History

REV#	DATE	AFFECTED SECTION	AUTHOR
1	2025/03/18	1. Limitations Focus 작성	김수영
2	2025/03/20	2. Limitations and Research Gaps 작성	김수영
3	2025/03/21	1. Survey Paper 작성	김수영
4	2025/03/21	검토 및 첨삭	송재현, 김수영

Table of Contents

1.	SURVEY PAPER - LIMITATIONS FOCUS.....	4
2.	LIMITATIONS AND RESEARCH GAPS.....	11

1. Survey Paper - Limitations Focus

설문조사 결과

연구주제에 대한 설문조사는 크게 두 가지 관점에 대한 설문을 실시하였고 각 관점 당 두 가지 문항을 넣어서 총 네 문항에 대한 설문을 실시하였다.

첫 번째 관점은 Multi-Task Learning에 대한 인식을 물어보는 것으로 Multi-Task Learning의 상용화 가능성과 기존에 알려져 있는 학습방법인 Single Task Learning에 비해 더 유용한 학습방법일 것이라 생각하는지를 물어보았다.

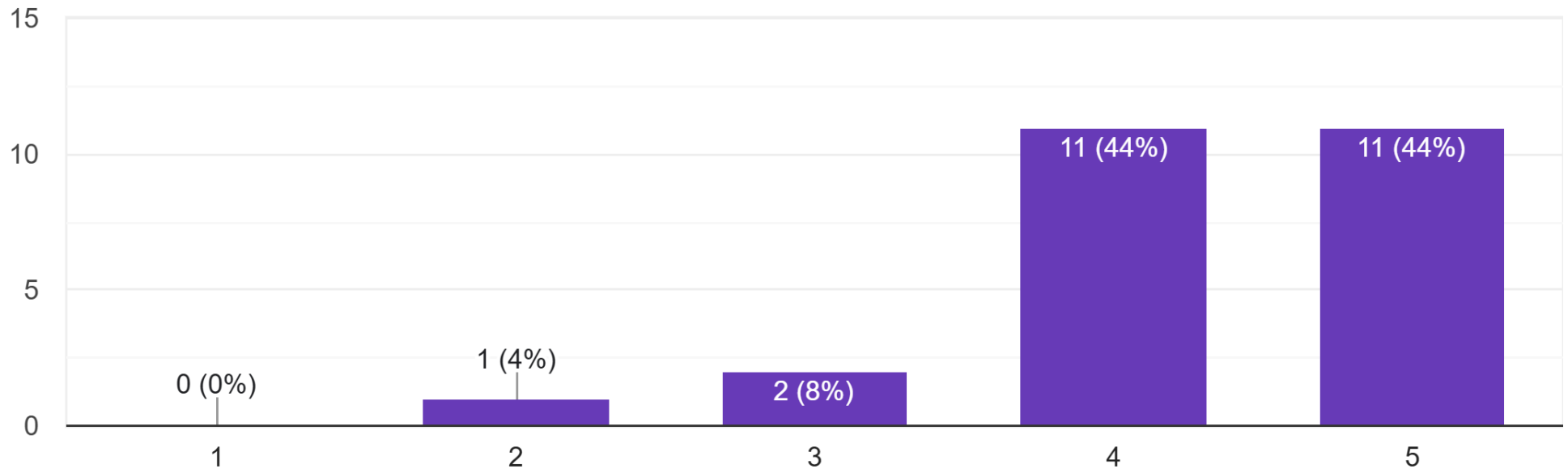
두 번째 관점은 Multi-Task Learning의 연구 방향 및 필요성 평가에 대한 것이다. Multi-Task Learning의 유효성 검증에 관한 연구의 필요성과 해당 Multi-Task Learning이 향후 AI 성능 발전에 필요한 기술인지 물어보았다.

조사 방법은 Goolge Form을 이용하여 온라인으로 진행하였고 조원의 지인이나, 학교 커뮤니티에 홍보하는 등의 방식으로 설문을 진행하였으며 총 25명이 응답해주었다. 엄밀하게 표집을 진행한 것은 아니나 과목과 과제의 특성 그리고 시간의 한계와 같은 여러 요인을 고려하면 이러한 방식이 그나마 단기간 내 많은 응답을 얻어낼 수 있는 방법이라 생각하였다. 25명이라는 응답 수치가 적다면 적은 수치라고 볼 수도 있으나 여건과 기간의 한계를 고려하면 이 정도의 수치만 되어도 설문을 진행한 가장 큰 목적인 대중들의 인식 파악은 달성한 수치라고 본다.

모든 문항들은 1점(매우 그렇지 않다)부터 5점(매우 그렇다)까지의 점수로 평가를 받도록 구성되어있으며 엄밀한 표본 조사 결과 보고서는 아니기에 Google Form에서 제공해주는 그래프를 간단하게 분석하는 방향으로 결과 분석을 수행하였다.

Multi-Task Learning을 활용한 AI 모델 개발이 상용화될 가능성이 있다고 생각하십니까?

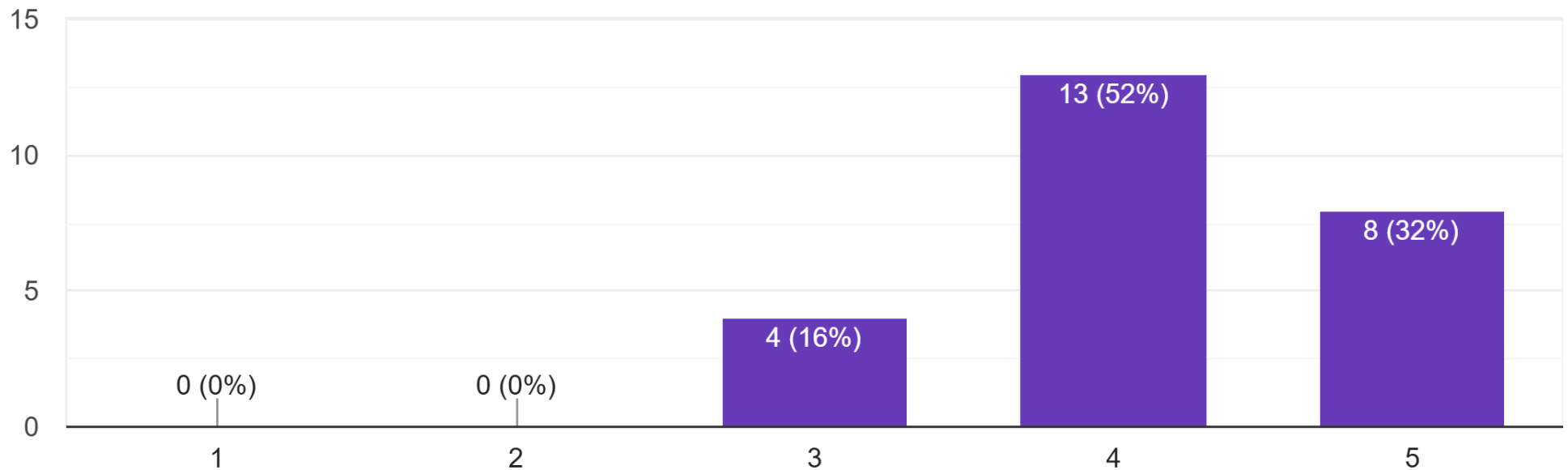
응답 25개



첫 번째 문항으로 Multi-Task Learning의 상용화에 관한 문항이다. 대부분의 사람들이 Multi-Task Learning의 상용화에 긍정적인 답변을 해주었다. 그렇다(4점) 이상인 반응이 전체 응답 중 88%를 차지하고 있으며 부정적인 응답은 단 한 건만 존재하였다. 이는 Multi-Task Learning이라는 방법론을 활용한 모형의 상용화 가능성에 대한 밝은 전망을 나타낸다고 볼 수 있다. 또한, 대중들이 Multi-Task Learning이라는 학습 방법이 적용된 모형에 대해 대중들이 거부감 없이 수용할 준비가 되어 있으며, 실제 서비스로 접했을 때 활용 의사 또한 높을 수 있음을 시사한다.

Multi-Task Learning이 단일 작업 학습(Single Task Learning, 기존에 사용되는 학습 방법)보다 유용할 가능성이 있다고 생각하십니까?

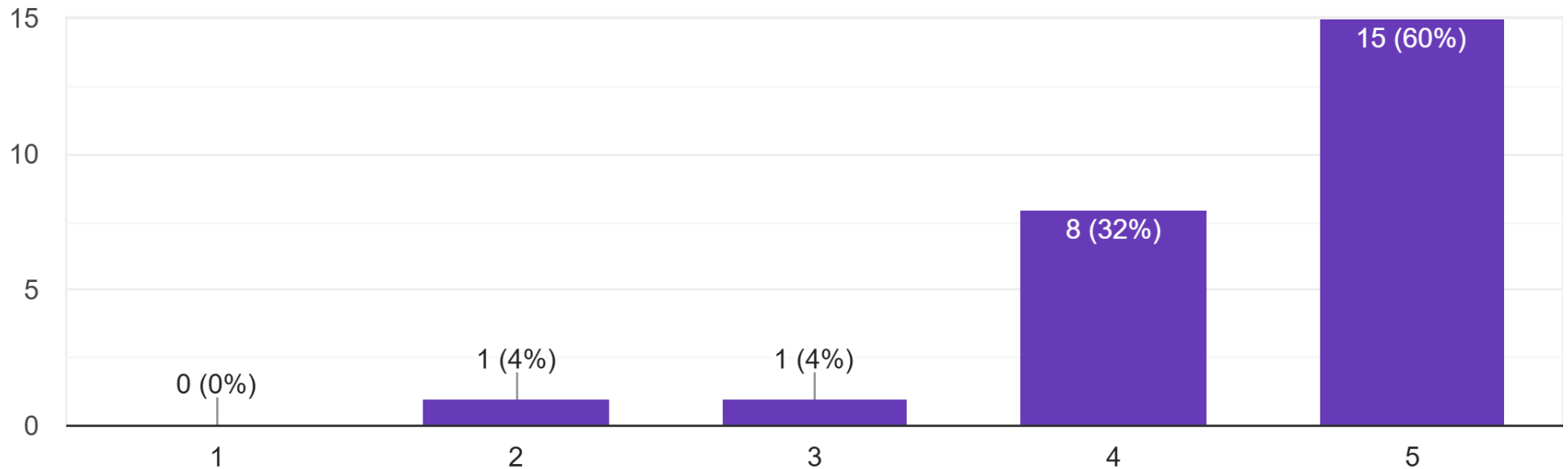
응답 25개



두 번째 문항은 Multi-Task Learning이 단일 작업 학습(Single Task Learning)보다 유용하다고 생각하는지에 대한 문항이다. ‘그렇다(4점)’ 이상의 응답 비율이 84%, ‘보통이다(3점)’ 이상의 응답 비율이 100%로, 부정적인 응답은 전혀 없었다. 이는 Multi-Task Learning에 대한 대중의 인식이 전반적으로 긍정적임을 의미하며, 앞선 상용화 관련 응답과 연계해보았을 때, Multi-Task Learning이 Single Task Learning보다 더 높은 성능을 낼 것이라는 기대가 상용화에 대한 긍정적 평가로 이어졌음을 유추할 수 있다. 결과적으로, Multi-Task Learning은 대중에게 기대를 받고 있는 학습 방식이며, 현실적인 활용 가치가 높은 기술로 인식되고 있음을 알 수 있다.

Multi-Task Learning의 유효성 검증에 관한 연구가 필요하다고 생각하십니까?

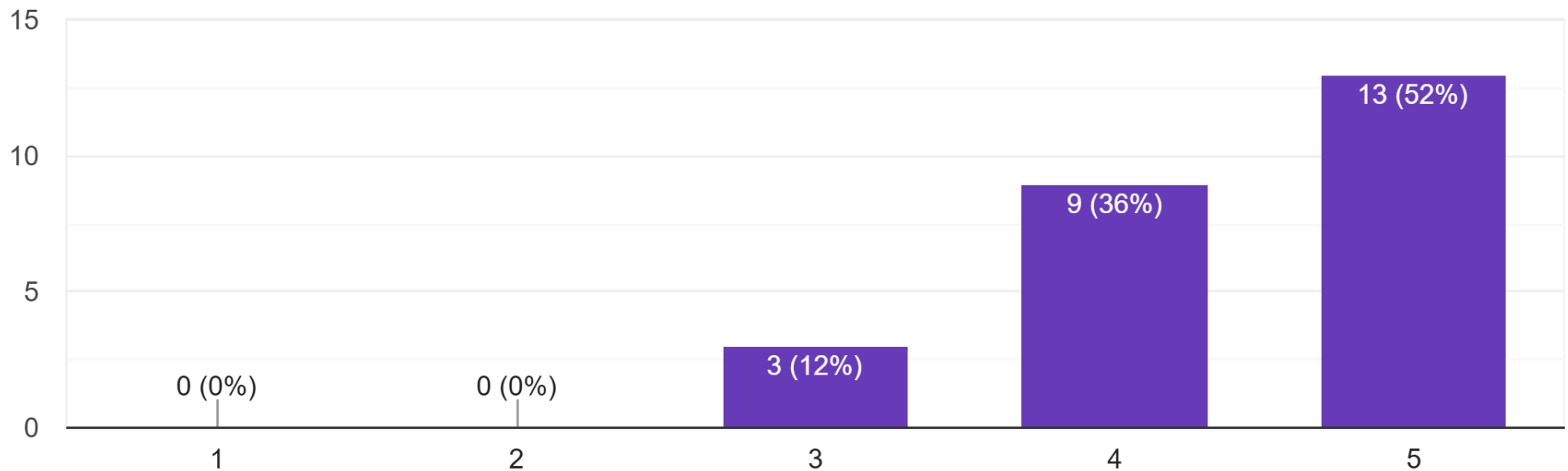
응답 25개



세 번째 문항은 Multi-Task Learning의 유효성 검증 필요성을 묻는 항목이다. 응답자의 60%가 ‘매우 그렇다(5점)’라고 응답하였으며, ‘그렇다(4점)’ 이상 응답은 전체 비율 중 92%를 차지한다. 이는 Multi-Task Learning의 효과에 대한 검증이 필요하다는 데 대중들이 높은 공감대를 형성하고 있음을 보여준다. 특히, 이전 문항에서 파악한 Multi-Task Learning이 Single Task Learning보다 나은 성능을 낼 것이라는 기대와 맞물려, MTL의 유효성을 실증적으로 검증하는 과정이 매우 중요하다고 인식됨을 시사한다. 이러한 결과는 우리 연구의 핵심 주제인 “Multi-Task Learning의 유효성 검증”이 시의적절하며 학문적·산업적 가치를 동시에 가질 수 있는 주제임을 뒷받침한다.

Multi-Task Learning 연구가 향후 AI의 성능 발전에 필요한 기술이라고 생각하십니까?

응답 25개



마지막 문항은 Multi-Task Learning 연구가 AI 성능 발전에 필요한 기술인지를 묻는 항목이다. 역시나 ‘그렇다(4점)’ 이상인 비율이 88%를 차지하며 부정적인 응답은 단 한 건도 없음을 확인할 수 있다. 이는 Multi-Task Learning이 단순한 이론적 개념을 넘어, AI 기술의 실질적인 발전에 기여할 수 있는 핵심 기술로 인식되고 있음을 보여준다. 그렇기에 이러한 기대에 부응하기 위해서는 Multi-Task Learning의 효과를 입증할 수 있는 유효성 검증에 관한 연구도 당연히 수반되어야 함을 시사하며, 우리 연구 주제의 방향성과 필요성에 필요한 당위성을 더욱 뒷받침해준다. 마지막으로 이러한 결과는 Multi-Task Learning을 주제로 한 본 연구가 기술적 유의미성과 사회적 수요를 모두 충족시키는 주제임을 다시금 확인시켜주는 계기가 되었다.

종합설계 1

번호	연구 제목(저자)	저널/컨퍼런스 (연도)	주요 내용 요약	한계점
1	Learning Sparse Sharing Architectures for Multiple Tasks - Tianxiang Sun, Yunfan Shao, Xiaonan Li, Pengfei Liu, Hang Yan, Xipeng Qiu, Xuanjing Huang	AAAI/2020	Multi-Task Learning의 학습 방법론으로 알려진 대표적인 방법은 Hard Sharing, Soft Sharing, Hierarchical Sharing이 있다. 본 논문의 연구자들은 이러한 방법론들의 단점을 해소하기 위해 보다 일반화된 학습 방법을 고안하였다. Multi-Task Learning을 수행할 시 sparsity라는 값을 임계값으로 설정하여 Task 별로 특정 parameter만을 학습에 이용하는 방법이다. 이러한 방법을 연구자들은 Sparse Sharing이라 칭하였으며 이 방법은 보다 유연한 Hard Sharing이라 볼 수 있다. 특히, 기존 방식들은 작업 간의 관계를 이해하기 어렵고, 이질적인 작업들을 효과적으로 처리하거나 파라미터 효율성을 높이는 데 어려움을 겪는다는 점을 강조하였으며 Sparse Sharing은 이러한 문제점에 대해 효과적으로 대처할 수 있다 주장하고 있다. 이 방법은 뛰어난 경량화 정도를 보여주며 성능 역시 향상됨을 실험을 통해 입증하였다.	연구에서 실험에 사용한 Mask의 선정 과정은 단순히 각 Task 별로 Single Task Learning을 수행할 때 성능이 가장 좋은 Mask를 선택하는 방식으로 하였다. 지식 전이와 공유 지식을 통해 학습한다는 Multi-Task Learning 철학에는 다소 어긋난 형식의 방안이라 생각한다. 또한, 이러한 선정 방식은 시간 역시 굉장히 오래 걸린다. sparsity라는 지표를 threshold로 설정한 후 sparsity가 낮은 parameter들을 일정 비율만큼 없애주는 역할을 하는 것이 Mask인데 Mask가 적용된 parameter로 처음부터 재학습시키는 방식으로 모형을 구하고 그 모형에 대한 STL 성능평가를 한 뒤 이 결과들을 바탕으로 Mask를 선정한다. 이는 굉장히 비효율적이며 효율적으로 Mask를 구하는 방안을 탐색하는 것이 필요하다 생각한다.
2	Real-World Image Super-Resolution as Multi-Task Learning - Wenlong Zhang, Xiaohui Li, Guangyuan Shi, Xiangyu Chen, Xiaoyun Zhang, Yu Qiao, Xiao-MingWu, Chao	NeurIPS/2023	실제 세계 이미지 초해상도(real-SR) 문제를 Multi-Task Learning 관점에서 분석하고, 작업 간 경쟁(Task Competition) 문제를 해결하기 위한 새로운 학습 방법을 제안하였다. 기존 real-SR 방식은 여러 개의 열화(degradation) 작업을 단일 모델에서 처리하는 방식이었으며, 이로 인해 특정 작업이 우세하게 학습되면서 다른 작업의 성능이 저하되는 문제가 발생했다. 연구자	작업 그룹화의 기준이 성능 향상 점수(Performance Improvement Score, PIS)를 기반으로 설정되었는데, 이는 데이터셋과 모형 구조에 따라 최적의 임계값이 달라질 수 있으며, 보편적으로 적용하기 어렵다는 문제가 있다. 또한, 작업 그룹을 활용한 Fine Tuning 과정에서도 문제가 발생할 수 있다. 특정 작업 그룹이 학습 과정에서 다른 작업보다 더 많이 고려될 경

종합설계 1

	Dong		들은 이를 해결하기 위해 Task Grouping-based Real-SR(TGSR) 방법을 개발하였다. 성능이 낮은 작업을 식별한 후 여러 개의 작업 그룹(Task Groups)으로 나누고, 이를 활용해 미세 조정(Fine-tuning)을 수행하는 방식으로 문제를 해결하였다. 이를 통해 작업 간 균형을 유지하고, 전체 모델 성능을 향상시킬 수 있음을 실험을 통해 입증하였다..	우, 일부 작업의 학습이 여전히 제한될 가능성이 있다. 또한, 그룹화 과정에서 성능이 낮은 작업을 식별하는 데 시간이 많이 소요될 수 있으며, 연산 비용이 증가하는 문제가 존재한다. 마지막으로 실험에서는 PSNR과 LPIPS 같은 정량적 성능 평가를 중심으로 검증이 이루어졌지만, 사람이 직접 평가하는 시각적 품질(Visual Perception)에 대한 연구는 부족했다. 모델이 생성한 이미지의 실제적인 시각적 품질이 향상되었는지에 대한 추가적인 검증이 필요하다.
3	Efficiently Identifying Task Groupings for Multi-Task Learning - Christopher Fifty, Ehsan Amid, Zhe Zhao, Tianhe Yu1, Rohan Anil1, Chelsea Finn1	NeurLPS/2021	기존 Multi-Task Learning에서 작업(Task)를 선정하는 방법은 모든 작업을 하나의 모델에서 학습하거나 가능한 모든 작업 간 조합을 탐색하는 방식이 주로 사용되었다. 하지만 전자는 작업 간 간섭으로 인해 성능 저하를 유발할 수 있고, 후자는 연산 비용이 지나치게 높다는 문제가 있다. 본 논문에서는 작업 간 그래디언트 상호작용을 분석하여 작업 간 유사도(Affinity)를 측정하고, 이를 바탕으로 최적의 태스크 그룹을 자동으로 구성하는 Task Affinity Grouping(TAG) 프레임워크를 제안하였다. TAG의 핵심 원리는 하나의 작업의 그래디언트 업데이트가 다른 태스크의 손실에 미치는 영향을 정량화하여, 서로 긍정적인 영향을 주는 작업들을 함께 그룹화하는 것이다. 이 방법은 기존의 Task 선정 방식보다 월등히 빠른 속도로 Task Group을 구성하였으며 성능 역시 향상됨을 보였다. 또한, TAG 그룹핑을 기반	논문에서 제시된 이론적 분석은 convex loss function(볼록 손실 함수)를 바탕으로 제한한 설정에서 이루어졌으며, 실제 딥러닝 모델의 비선형적이고 non-convex한 손실 함수에 대한 이론적 분석은 이루어지지 않았다. 또한, 논문에서는 세 개 이상의 작업들로 구성된 그룹의 유사도(Affinity)를 단순히 해당 그룹에서의 쌍이 되는 조합들의 유사도를 모두 구한 후 그 평균을 해 해당 그룹의 유사도로 근사하는 방식으로 설정하였는데 보다 통계적으로 유의미한 그룹핑 전략이 필요하다 생각된다. 또한, 논문 내에서 자체적으로 지정한 내용으로 Task Affinity를 단순히 “작업 간 유사성”을 나타내는 척도로 오해해버려 인과 관계가 없는 결론을 도출해버릴 위험이 있다. 즉, Task Affinity에 대한 해석이 변질될 우려가 있다는 것이다. 이는 민감한 정보 (ex 의학 데이

		으로 기존에 알려진 최적화 기법(PCGrad 등)을 추가로 적용했을 때 추가적인 성능 향상도 확인되었다.	터, 법률 데이터)를 다루는 데이터셋에는 보다 신중하게 적용시켜야 한다는 한계가 있다.
--	--	--	--

2. Limitations and Research Gaps

번호	기존 연구	한계점	연구 필요성	본 연구의 기여
1	Learning Sparse Sharing Architectures for Multiple Tasks - Tianxiang Sun, Yunfan Shao, Xiaonan Li, Pengfei Liu, Hang Yan, Xipeng Qiu, Xuanjing Huang	논문에서는 각 태스크별 Single Task Learning(STL)을 수행한 후 성능이 가장 좋은 Mask를 선택하는 방식을 사용했다. 그러나 이 방식은 Multi-Task Learning(MTL)의 핵심 개념인 지식 공유와 전이 학습의 원리에 맞지 않으며, 작업 간의 유용한 공유 구조를 충분히 반영하지 못하는 한계가 있다. 또한, 실험에서 적절한 Mask를 얻기까지의 학습 과정이 비효율적이고 시간이 과도하게 소요된다. 이러한 문제를 해결하기 위해 보다 효율적인 Mask 선정 방법이 필요하다.	Sparse Sharing 구조는 정적인 공유구조이다. 사전에 작업 별로 학습일 진행하여 얻은 Mask를 갖고 이를 Multi-Task Learning에 활용하는 방식인 것이다. 따라서, Mask를 구하고 이를 적용하는 과정이 보다 동적으로 진행되어 공유 구조를 조절할 수 있는 메커니즘 내지 알고리즘을 개발하는 것이 필요하다 생각한다. 또한, 이러한 개발에는 한계점에서도 언급하였던 효율적인 Mask를 구하는 방안과 연계하여 Mask를 빠르게 구할 수 있게 하는 최적화 연구 역시 동반되어야 한다.	해당 연구는 새로운 파라미터 공유 메커니즘인 Sparse Sharing을 제시하였다. 기존에 알려져 있는 방식인 Hard Sharing, Hierarchical Sharing을 더욱 유연하게 다듬어 학습 방법의 일반성을 높였다. 이를 통해 부정적 전이(Negative transfer)를 완화시키는 효과를 제시하였고 이는 모형의 일반화를 내세우는 Multi-Task Learning의 지론에 걸맞은 방법임을 보여주었다. 즉, 보다 일반화된 새로운 학습 방법을 제시하며 이에 대한 성능과 경량화의 정도를 실험적으로 증명하여 기존 Multi-Task Learning의 주요 학습 방법의 단점들을 충분히 보완할 수 있는 것이 가능성을 시사한다.
2	Real-World Image Super-Resolution as	작업 그룹화 기준이 성능 향상 점수를 기반으로 설정되었지만, 이	Task Grouping-based Real-SR(TGSR) 방법은 성능이 낮은 작업들을 Grouping한	해당 연구는 Real-SR이라는 문제를 Multi-Task Learning 문제로 정의한 것에 큰 의미가 있다.

종합설계 1

	<p>Multi-Task Learning - Wenlong Zhang, Xiaohui Li, Guangyuan Shil, Xiangyu Chen, Xiaoyun Zhang, Yu Qiao, Xiao-MingWu1, Chao Dong</p>	<p>는 데이터셋과 모델 구조에 따라 최적의 임계값이 달라져 보편적인 적용이 어렵다. 또한, Fine-Tuning 과정에서 특정 작업 그룹이 과도하게 학습될 경우 일부 작업의 학습이 제한될 가능성이 있으며, 성능이 낮은 작업을 식별하는 과정에서 연산 비용이 증가하는 문제가 있다. 실험에서는 정량적 성능 평가에 집중되었으나, 시각적 품질에 대한 추가 검증이 부족하여 모델이 생성한 이미지의 실제 품질 향상을 확인할 필요가 있다.</p>	<p>다. 즉, 사전에 정적으로 Grouping을 수행하기에 추후 학습이 진행됨에 따라 작업 간 관계가 변화될 가능성이 배제된다. 또한, 한계점에서도 나왔듯이 데이터셋과 모델의 구조에 따라 작업 그룹화의 기준 임계점이 달라짐을 파악하였으므로 이러한 문제를 해결하기 위해서라도 따라서 동적으로 작업 그룹을 재구성하는 메커니즘에 대한 연구가 필요하다 생각한다. 이러한 연구는 Fine-Tuning 과정에서 특정 작업 그룹이 과도하게 학습되는 것 역시 미연에 방지할 수 있기에 연구 의의가 크다고 생각한다.</p>	<p>연구자들은 상기한 문제를 해결하기 위해 Multi-Task Learning에 본인들이 고안한 알고리즘을 적용시켜 실험적으로 검증해냈다. 이는 Multi-Task Learning이 비단 AI 모형의 경량화 및 성능 개선 문제에만 초점을 맞춘 것뿐만 아니라 저해상도 이미지를 고해상도 이미지로 복원하는 문제(Real-SR)처럼 우리가 해결하고자 하는 다른 문제들에도 충분히 적용할 수 있음을 여실히 보여준다. 즉, Multi-Task Learning의 주요 목표인 “일반화”가 단순히 학습 모형의 구조에만 국한된 것이 아닌 문제 해결이라는 더 큰 카테고리로 확장할 수 있음을 시사한다.</p>
3	<p>Efficiently Identifying Task Groupings for Multi-Task Learning - Christopher Fifty, Ehsan Amid, Zhe Zhao, Tianhe Yu1, Rohan Anil1, Chelsea Finn1</p>	<p>논문의 이론적 분석은 convex loss function을 기반으로 제한된 설정에서 이루어졌으며, 실제 딥러닝 모델의 비선형적이고 non-convex한 손실 함수에 대한 이론적 분석은 포함되지 않았다. 또한, 세 개의 이상의 작업 그룹에서 유사도를 구하는 방식에는 보다 통계적으로 정교한 그룹핑 전략이 필요하다. 더불어, Task Affinity를 단순히 "작업 간 유사성"으로 해석하면,</p>	<p>본 연구에서는 세 개 이상의 작업이 있는 작업 그룹의 유사도는 단순히 작업 간 쌍의 조합들의 평균으로 정의를 하였다. 이에 대해 보다 통계적 유효성이 검증된 방식으로 그룹의 유사도를 구한다거나 (유사도 간 T-Test 혹은 F-Test 수행, Regularization 등) 합리적인 메커니즘을 개발하는 연구가 필요하다 생각한다. non-convex한 손실 함수에 대한 이론적 분석 및 증명 역시 추후에 검증돼야할 부분이라 생각한다.</p>	<p>해당 연구는 Multi-Task Learning에서 학습시킬 작업들을 어떻게 하면 효율적이고 객관적으로 선택할 수 있는지를 제시하였다. 그리고 이에 대한 이론적 증명 및 실험적 검증을 수행하여 그 효용성을 입증하였다. 특히, 기존의 Single Task Learning이나 모든 작업을 학습시키는 Multi-Task Learning(Hard sharing) 보다 우수한 성능을 보여주었다. 즉, Multi-Task Learning의 주된 목표인 일반화에 대한 객관적인 지표(Affinity)를 마련해주었고, 경량화 역시 달성하였다. 이는 Multi-Task Learning이</p>

종합설계 1

		인과 관계가 없는 잘못된 결론을 도출할 위험이 있다. 특히, 의료 및 법률 데이터와 같은 민감한 정보를 다루는 환경에서는 적용이 어렵다는 한계가 존재한다.	마지막으로, 민감한 데이터셋을 활용할 경우 유사도에 대한 기준 및 해석 방안을 타 분야의 전문가와 의논하여 가이드라인을 제시하는 등의 조치를 하는 식의 방안으로 보완할 수 있을 것이다.	추구하는 목표를 동시에 달성할 수 있는 방안이 존재함을 시사하며 작업의 선정 역시도 학습 과정에 대단히 중요한 측면을 담당함을 함의한다 볼 수 있다.
--	--	--	---	---