

# 검색공간의 점진적인 축소를 통한 DASH 성능 향상에 관한 연구

이정은, 이정우  
서울대학교

j\_lee@snu.ac.kr, junglee@snu.ac.kr

## A study on improving DASH performance through progressive reduction of search space

Lee Jung Eun, Lee Jung Woo  
Seoul National Univ.

### 요약

본 논문은 다양한 분야의 문제를 해결하기 위해 도입된 DASH(Diverse-task Architecture Search)로 생성된 신경망의 구조가 다양하면서 구조마다 성능의 편차가 있음을 확인하고 DASH의 집합된 컨볼루션의 검색공간을 점진적으로 축소하여 신경망의 성능을 개선하고 그 편차를 감소시켰다.

### I. 서론

신경망 아키텍처 검색(NAS)은 심층 신경망 설계를 간소화하는 것을 목표로 모델 개발을 자동화하는 중추적인 도구로 등장했다. NAS는 복잡한 아키텍처를 튜닝할 때 소요되는 인력을 최소화하면서 전문가들이 수년간 설계한 아키텍처에 필적하는 성능을 보장한다. 대표적으로 유명한 NAS 방식인 DARTS(Differentiable Architecture Search)[1]는 아키텍처 검색 프로세스에 차별화를 도입하여 경사 하강(gradient) 기반 최적화 방법을 사용하고 개별 아키텍처 검색을 지속적인 최적화 문제로 변환했다. 또한 Progressive DARTS(P-DARTS)[2]같이 DARTS의 학습 안정성을 높여 아키텍처의 성능을 높이는 방식도 등장했다. 그러나 다양한 NAS 평가 방법을 강조하는 최신 벤치마크인 NAS-Bench-360[3]로 분석한 결과, DARTS를 포함한 NAS 방법은 원래의 목적(DARTS의 경우, CIFAR와 ImageNet이 주 목적이었다.)을 벗어난 문제에 사용될 때 좋은 성능이 보장되지 않는다는 것이 밝혀졌다. 이를 해결하기 위해 검색 속도와 효율성을 유지하면서 높은 표현력을 통해 도메인 전반에 걸쳐 높은 정확도를 나타내는 아키텍처를 검색하는 DASH(Diverse-task Architecture Search)[4]가 등장했다. DASH의 표현력은 검색공간의 크기 의해 결정되고 정확도는 검색공간의 적합성 의해 결정되는데 검색 효율성과 최종 아키텍처의 성능을 위해서는 효율적인 검색공간을 구성하는 것이 중요하다. 본 연구에서는 기존 DASH의 검색공간으로부터 검색과정을 반복하고 이 반복과정에서 점진적으로 검색공간을 축소시켜 기존 방식과의 성능을 비교하였다.

### II. 본론

DASH는 CNN을 백본(backbone) 네트워크로 사용하여 CNN 백본에 존재하는 각각의 컨볼루션 연산을 검색공간이 집합된(aggregated) 컨볼루션으로 대체한 후 학습을 진행하는데 검색과정에서 문제에 적합한 커널 크기와 확장(dilation)이 결정된다. aggregated 컨볼루션이 제공하는 표현력으로 NAS-Bench-360의 10가지 작업에 대해서 전반적으로 좋은 성능을 나타낼 수 있었다. 그러나 우리는 검색공간 설정이 아키텍처 성능에 영향을 미친다는 점에 주목했다. 본 논문에서 우리는 P-DARTS에서 사용된 방식을 참고하였고, 이에 따라 커널 패턴 검색을 두 차례 이상 반복하여 검색공간을 점진적으로(Progressive) 축소하는 실험을 진행하였다. 실험을 위해 NAS-Bench-360의 10가지 작업 중 3가지를 선정하였고, 각 작업에 대한 정보는 표 1에서 확인할 수 있다.

표 1. 실험을 위한 NAS-Bench-360의 3가지 작업.  
커널 크기와 확장에 의해 검색공간이 정해진다.

작업	차원	Type	커널 크기	확장
NinaPro	2D	Point	[3, 5, 7, 9]	[1, 3, 7, 15]
DarcyFlow	2D	Dense	[3, 5, 7, 9]	[1, 3, 7, 15]
ECG	1D	Point	[3, 7, 11, 15, 19]	[1, 3, 7, 15]

백본에서 각 레이어는 (커널 크기 수  $\times$  확장 수)만큼의 커널 패턴이 집합된 aggregated 컨볼루션으로 대체되어 학습이 진행된다. 본 논문의 실험에서는 첫 번째 커널 패턴 검색 시 커널 패턴의 순위를 매겨 레이어마다 가능성이 낮은 커널 패턴을 제거하였고, 이렇게 만들어진 새로운 aggregated 컨볼루션으로 각 레이어를 다시 대체하여 두 번째 커널 패턴을 검색하여 각각의 레이어 별 커널 패턴을 확정하였다. 한편, 2D

작업에서 백본 네트워크로 사용된 WRN 16-1 과 WRN 16-4 에는 13 개의 컨볼루션 레이어가 존재하고 1D 테스트에서 사용된 1D WRN 에는 9 개의 컨볼루션 레이어가 존재한다. 즉, 첫 번째 검색 시에는 2D 에서는 13 개 레이어 모두에, 1D 에서는 9 개 레이어 모두에 같은 aggregated 컨볼루션이 존재하지만 두 번째 검색부터 각 레이어 마다 검색공간이 달라진다. 본 논문에서는 표현의 편의를 위해 이 방식을 Progressive DASH, P-DASH 로 명명하였다.

**표 2. DASH 와 P-DASH 의 작업 별 에러율.**  
에러율은 실험을 3 회 진행한 결과에 대한 평균이다.

	NinaPro error (%)	Darcy Flow Relative l2	ECG 1 - F1
DASH	6.60±0.33	0.0079±0.002	0.32±0.007
P-DASH	5.90±0.08	0.0063±0.0008	0.30±0.001

정확한 분석을 위해 3 가지 작업에 대해 각각 3 회씩 실험을 진행하였고 결과는 표 2 에서 볼 수 있다. DASH 방식에 비해 Progressive DASH 가 NinaPro 에서 11%, Darcy Flow 에서 20%, 그리고 ECG 에서 7%의 에러율 감소를 보였고 평균 12.6% 감소했다. 또한 에러율의 편차도 DASH 의 실험결과와 비교하여 작아진 것을 볼 수 있다. 이는 검색공간을 점진적으로 축소하는 과정에서 작업에 필요 없는 커널 패턴이 제거되어 검색공간의 효율이 증가했기 때문이라고 볼 수 있다.

### III. 결론

점진적으로 DASH 의 검색공간을 축소해 나가는 방식은 기존의 DASH 방식보다 안정적으로 더 좋은 성능의 아키텍처 검색을 가능하게 했다. NAS-Bench-360 의 나머지 7 가지 작업에 대해서도 안정적으로 좋은 성능의 아키텍처 검색이 가능할 것으로 기대된다. 향후에는 배경지식이 없거나 알려지지 않은 도메인에 대해서 학습 시 검색공간을 임의로 설정하게 되어도 검색을 반복하면서 점진적으로 검색공간을 업데이트 하면 최종적으로는 좋은 아키텍처 검색이 가능할 것이다.

### ACKNOWLEDGMENT

This work is supported by Institute of Information & communications Technology Planning & Evaluation (IITP, 2021-0-00106) grant funded by the Ministry of Science and ICT (MSIT) INMAC, and BK21 FOUR program.

### 참 고 문 헌

[1] Hanxiao Liu, Karen Simonyan, and Yiming Yang. Darts: Differentiable architecture search. arXiv preprint arXiv:1806.09055, 2018.

[2] Xin Chen, Lingxi Xie, Jun Wu, and Qi Tian. Progressive differentiable architecture search: Bridging the depth gap between search and evaluation. In Proceedings of the IEEE/CVF international conference on computer vision, pages 1294- 1303, 2019

[3] Renbo Tu, Nicholas Roberts, Misha Khodak, Junhong Shen, Frederic Sala, and Ameet Talwalkar. Nas-bench-360: Benchmarking neural architecture search on diverse tasks. Advances in Neural Information Processing Systems, 35:12380-12394, 2022.

[4] Junhong Shen, Misha Khodak, and Ameet Talwalkar. Efficient architecture search for diverse tasks. Advances in Neural Information Processing Systems, 35:16151-16164, 2022.