The Journal of the Convergence on Culture Technology (JCCT)
Vol. 9, No. 1, pp.649-654, January 31, 2023. pISSN 2384-0358, eISSN 2384-0366

http://dx.doi.org/10.17703/JCCT.2023.9.1.649

JCCT 2023-1-79

# 그래프 합성곱-신경망 구조 탐색 : 그래프 합성곱 신경망을 이용한 신경망 구조 탐색

# Graph Convolutional - Network Architecture Search : Network architecture search Using Graph Convolution Neural Networks

최수연*, 박종열**

**Su-Youn Choi*, Jong-Youel Park****

**요 약** 본 논문은 그래프 합성곱 신경망을 이용한 신경망 구조 탐색 모델 설계를 제안한다. 딥 러닝은 블랙박스로 학습이 진행되는 특성으로 인해 설계한 모델이 최적화된 성능을 가지는 구조인지 검증하지 못하는 문제점이 존재한다. 신경망 구조 탐색 모델은 모델을 생성하는 순환 신경망과 생성된 네트워크인 합성곱 신경망으로 구성되어있다. 통상의 신경망 구조 탐색 모델은 순환신경망 계열을 사용하지만 우리는 본 논문에서 순환신경망 대신 그래프 합성곱 신경망을 사용하여 합성곱 신경망 모델을 생성하는 GC-NAS를 제안한다. 제안하는 GC-NAS는 Layer Extraction Block을 이용하여 Depth를 탐색하며 Hyper Parameter Prediction Block을 이용하여 Depth 정보를 기반으로 한 spatial, temporal 정보(hyper parameter)를 병렬적으로 탐색합니다. 따라서 Depth 정보를 반영하기 때문에 탐색 영역이 더 넓으며 Depth 정보와 병렬적 탐색을 진행함으로 모델의 탐색 영역의 목적성이 분명하기 때문에 GC-NAS대비 이론적 구조에 있어서 우위에 있다고 판단된다. GC-NAS는 그래프 합성곱 신경망 블록 및 그래프 생성 알고리즘을 통하여 기존 신경망 구조 탐색 모델에서 순환 신경망이 가지는 고차원 시간 축의 문제와 공간적 탐색의 범위 문제를 해결할 것으로 기대한다. 또한 우리는 본 논문이 제안하는 GC-NAS를 통하여 신경망 구조 탐색에 그래프 합성곱 신경망을 적용하는 연구가 활발히 이루어질 수 있는 계기가 될 수 있기를 기대한다.

**주요어** : 신경망 구조 탐색, 그래프 합성곱 신경망, 순환 신경망, 모델 구조 설계

**Abstract** This paper proposes the design of a neural network structure search model using graph convolutional neural networks. Deep learning has a problem of not being able to verify whether the designed model has a structure with optimized performance due to the nature of learning as a black box. The neural network structure search model is composed of a recurrent neural network that creates a model and a convolutional neural network that is the generated network. Conventional neural network structure search models use recurrent neural networks, but in this paper, we propose GC-NAS, which uses graph convolutional neural networks instead of recurrent neural networks to create convolutional neural network models. The proposed GC-NAS uses the Layer Extraction Block to explore depth, and the Hyper Parameter Prediction Block to explore spatial and temporal information (hyper parameters) based on depth information in parallel. Therefore, since the depth information is reflected, the search area is wider, and the purpose of the search area of the model is clear by conducting a parallel search with depth information, so it is judged to be superior in theoretical structure compared to GC-NAS. GC-NAS is expected to solve the problem of the high-dimensional time axis and the range of spatial search of recurrent neural networks in the existing neural network structure search model through the graph convolutional neural network block and graph generation algorithm. In addition, we hope that the GC-NAS proposed in this paper will serve as an opportunity for active research on the application of graph convolutional neural networks to neural network structure search.

**Key words :** Neural Architecture Search, Graph Convolutional Network, Recurrent Neural Network, Model Architecture Design.

*정회원, 숭실대학교 베어드학부 겸임교수 (제1저자)
(배화여자대학교 겸임교수, ㈜블루오션정보통신 선임연구원)
**정회원, 배화여자대학교 스마트IT학과 조교수 (교신저자)
접수일: 2022년 12월 26일, 수정완료일: 2023년 1월 5일
게재확정일: 2023년 1월 9일

Received: December 26, 2022 / Revised: January 5, 2023
Accepted: January 9, 2023
**Corresponding Author: pjy@baewha.ac.kr
Dept. of SmartIT, Baewha Women's Univ, Korea

## Ⅰ. 서 론

머신러닝(Machine learning, ML)과 딥 러닝(Deep learning, DL)은 다양한 분야에 사람보다 뛰어난 성능을 보이며 각광받고 있는 분야이다. 근래 인기에 힘입어 딥 러닝 모델은 다양한 분야에서 많은 연구가 이루어져 꾸준히 새로운 모델 구조가 개발되고 있다. 하지만 딥 러닝은 블랙박스로 학습이 진행되는 특성으로 인해 설계한 구조가 최고의 성능을 가진 합리적인 구조라는 근거를 제시하지 못한다. 이와 같은 문제로 수많은 연구자들은 직접 모델의 설계와 실험을 반복하며 최적의 모델 구조임을 증명하는 방법을 사용한다. 이러한 모델 구조의 개발은 전문적 지식을 요하며 많은 시간과 노력이 들어간다는 단점이 존재한다. 이에 대한 문제를 해결하고자 Barret Zoph은 신경망 구조 탐색(Network Architecture Search, NAS)이라는 새로운 모델 구조 개발 방법을 제안하였다[1]. 신경망 구조 탐색은 모델을 생성하는 순환 신경망(Recurrent Neural Network, RNN)과 생성된 신경망인 합성곱 신경망(Convolution Neural Network, CNN)이 존재한다[2][3][4]. 순환 신경망으로 합성곱 신경망 모델에 들어가는 레이어를 추출하고 추출한 레이어로 합성곱 신경망 모델을 구성하여 학습하는 방법을 통해 구성된 합성곱 신경망 모델의 성능을 평가한다. 여기서 순환 신경망 모델은 평가된 합성곱 신경망 모델의 성능 평가를 이용하여 합성곱 신경망이 더 높은 성능을 낼 수 있도록 학습하게 된다. 기본적인 신경망 구조 탐색은 순환 신경망 계열을 사용하지만 본 논문에서는 순환 신경망 대신 그래프 합성곱 신경망(Graph Convolutional Network, GCN)을 사용하여 합성곱 신경망 모델을 학습하는 그래프 합성곱 - 신경망 구조 탐색(Graph Convolutional - Network Architecture Search, GC-NAS)을 제안한다[5]. GC-NAS는 그래프 합성곱 모델을 도입하여 시계열적 관점에서 데이터의 상관관계를 분석하고 입력 값의 가중치를 종합적으로 검토하여 순환 신경망의 근본적인 문제점인 시계열 데이터의 입력적 한계의 문제와 각 입력 값의 가중치를 문제를 개선할 것으로 기대한다.

## Ⅱ. 이론고찰

신경망 구조 탐색은 순환 신경망을 사용하여 합성곱 신경망의 모델을 생성하고 생성된 합성곱 신경망 모델의 정확도를 최대화하기 위해 강화 학습으로 순환 신경망 모델을 훈련시켰다. 신경망 구조 탐색은 통상적으로 학습하는데 많은 시간과 비용이 든다[6].

딥 러닝 모델 구조를 최적화하는 연구는 이전부터 진행되어 왔다. 예를 들어 그림1(a)에 보인 바와 같이 전이 학습(Transfer Learning)은 이전에 대규모 데이터 셋을 학습한 딥 러닝 학습 모델의 구조와 가중치를 그대로 가져와 마지막 단계의 레이어만 수정하여 학습하는 방법으로 모델 설계의 시간을 크게 단축 시켰으며 소규모 데이터 셋을 학습하더라도 좋은 성능을 도출하는 성과를 보였다[7][8]. 그림1(b)에 보인 바와 같이 지식 증류(Knowledge Distillation)는 학습을 마친 큰 규모의 딥 러닝 학습 모델이 작은 규모의 딥 러닝 학습 모델에게 학습을 시키는 방법으로 불필요한 레이어와 가중치를 제거함으로써 성능은 비슷하게 유지하면서 더 작고 가벼운 딥 러닝 모델 구조를 만들 수 있게 하였다[9]. 하지만 위의 두 방법은 성능이 좋고 효율적인 최적의 딥 러닝 모델을 설계하는 방법이 아니며 여전히 연구자들의 끊임없는 실험을 통한 하이퍼 파라미터의 조정이 필요하다는 단점이 존재한다.

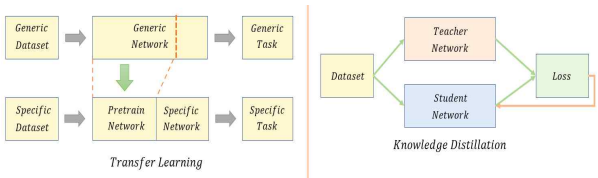


그림 1(a). 전이학습, 그림 1(b).지식증류
Figure 1(a). Transfer learning, Figure 1(b).Knowledge distillation

## Ⅲ. 연구방법

사람이 아닌 학습을 통해 만들어지는 딥 러닝 모델 구조가 정말 좋은 성능을 낼 수 있는지에 대한 검증은 다양하게 시도되어왔다. Efficient Net은 일부 레이어에 신경망 구조 탐색을 적용해 깊이, 넓이, 연결을 최적화시키는 레이어 구조를 설계함으로 모델 구조의 크기 대비 효과적인 성능을 달성하였다[10]. 또한 구글(Google)은 신경망 구조 탐색을 응용하여 반도체 회로를 설계하여 고성능의 TPU v4을 제작하였다[11]. 신경망 구조 탐색과 같은 모델을 설계 딥 러닝은 사람과 달리 직관적이지 못한 모델 구조를 만들어 내는 공통적인 특징이 있다.

이는 딥 러닝의 학습 과정이 사람과 같이 계층적, 점진적 학습을 하지 않는다는 것을 보여주며 반증적으로 사람의 논리, 직관적인 사고로는 최고의 모델을 설계할 수 없음을 보여준다. 이와 같은 이유로 신경망 구조 탐색은 최신 연구에 있어서 많은 관심을 받고 있다.

그래프 합성곱 신경망은 레이어간의 연결되는 관계를 파악하여 그래프를 생성 및 학습하는 딥 러닝 모델이다. 이론적으로는 그래프를 합성곱하여 그래프 간 상관 관계를 추론하고 가중치를 곱함으로 원하는 특징의 추출과 합성을 가능하게 한다. 하지만 그래프 합성곱 신경망을 시계열적 관점에서 보았을 때 이전 그래프와 현재 그래프의 상관 관계를 잘 추출하고 합성하기 때문에 현재 시점에도 이전의 연산된 데이터가 잘 반영될 수 있음을 알 수 있다. 이와 같은 특성으로 인해 그래프 합성곱 신경망은 순환 신경망을 대체 할 수 있다. 시공간 - 신경망 구조 탐색(Spatial Temporal - Graph Convolutional Network, ST-GCN)은 사람의 관절 좌표 데이터를 이용한 행동 인식 분야에서 전통적인 순환 신경망 계열의 딥 러닝 학습 모델을 제치고 최신기술(State Of The Art, SOTA)의 성능을 달성하였다[12]. ST-GCN은 사람의 관절 좌표를 시간과 공간축으로 배열한 뒤 그래프 합성곱 신경망을 통하여 시간, 공간 축으로 각각 그래프 연산을 적용하였다. 그래프 합성곱 신경망의 특성을 이용하여 현재 프레임의 관절 데이터는 이전 프레임의 주변 관절 데이터를 포함하여 계산하는 방식을 사용해 시공간 정보를 더욱이 세밀히 조합하여 높은 성능을 달성하였다.

신경망 구조 탐색 또한 순환 신경망 계열의 딥 러닝 모델 구조를 사용하기 때문에 그래프 합성곱 신경망을 적용한 연구가 있었다. 그림2(a)에 보인 바와 같이 그래프 합성곱 신경망 - 신경망 구조 탐색(Graph Convolution Network - Network Architecture Search, GCN-NAS)는 고차원적인 시간 축에 대응하기 어려운 그래프 합성곱 신경망의 단점을 고려하여 데이터를 시공간 축으로 모두 펼친 후 다중 동적 그래프 모듈(Multiple Dynamic Graph Module, MDGM)을 만들어 기존 그래프 합성곱 신경망의 한계점을 개선하여 적용한 신경망 구조 탐색을 소개하였다[13].

본 논문은 GCN-NAS와는 다른 별개의 구조로 GC-NAS를 제안한다. 그림2(b)에 보인 바와 같이 GC-NAS는 그래프 합성곱 신경망의 근본적인 특성을 이용한 신경망 구조 탐색으로 학습 과정 중 그래프 합성곱 신경망에서 추정한 레이어를 다음 레이어의 그래프 합성곱 신경망의 노드로 사용하여 기존 그래프 합성곱 신경망에서의 시간 축 문제를 해결 하였으며 사용한 레이어가 같은 공간 단의 노드로 존재함으로 공간적 정보의 상관 관계 또한 잘 파악할 수 있을 것으로 기대한다.

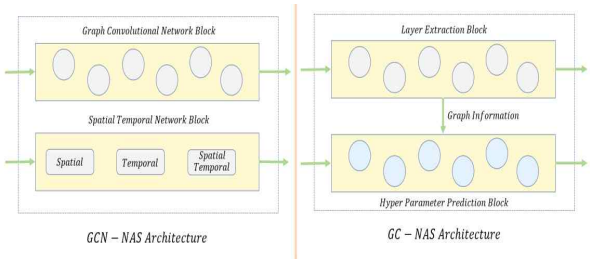


그림 2(a). GCN-NAS, 그림 2(b) GC-NAS 구조
Figure 2(a). GCN-NAS, Figure 2(b). GC-NAS structure

또한 그림3에 보인 바와 같이 GC-NAS는 기존 GCN-NAS 대비 구조가 간단하며 레이어가 쌓일수록 그래프에서 얻는 특징 정보가 정교해지기 때문에 GCN-NAS대비 좋은 성능을 달성할 것으로 기대한다.

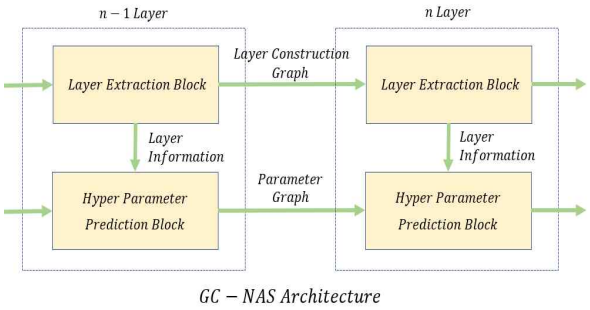


그림 3. GC-NAS의 전체 구조
Figure 3. Overall structure of GC-NAS

## Ⅳ. 연구결과

본 논문에서 제안하는 GC-NAS는 그래프 합성곱 신경망의 모델 구조를 기반으로 한다. 모델 생성은 그래프 합성곱 신경망, 생성된 모델은 합성곱 신경망으로 설계하였다. GC-NAS에서 사용하는 그래프 합성곱 신경망은 레이어를 도출하는 그래프 합성곱 신경망 블록과 하이퍼 파라미터를 추정하는 그래프 합성곱 신경망 블록으로 나뉜다. 레이어 도출 그래프 합성곱 신경망

블록은 그래프 합성곱 신경망의 n-1 레이어에서 출력으로 나온 레이어의 정보를 n번째 레이어 도출 그래프 합성곱 신경망 블록이 사용할 그래프 노드에 새로운 노드로 추가하여 새로운 레이어 도출 그래프 합성곱 신경망 블록 그래프를 생성한다. 여기서 사용하는 레이어 정보는 맵핑되어 있는 숫자로 conv2d, maxpool 등이 숫자와 일대일 매칭으로 되어있다. 앞선 방법으로 구성된 n번째 레이어의 모델 구조를 기반으로 하이퍼 파라미터 추정 그래프 합성곱 신경망 블록은 n번째 레이어의 파라미터와 연결을 재구성하게 된다.

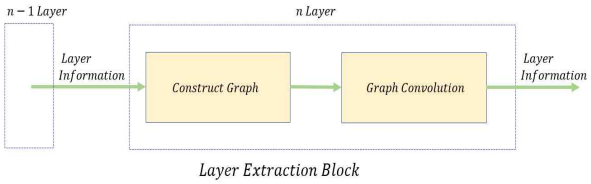


그림 4. GC-NAS의 레이어 추출 블록
Figure 4. Layer extraction block of GC-NAS

그림 4에 보인 바와 같이 레이어 도출 그래프 합성곱 신경망 블록은 이전 레이어 도출 블록이 추정한 레이어를 이용하여 현재까지 구성된 레이어 그래프를 생성하고 현재까지 구축된 합성곱 신경망의 그래프 정보를 기반으로 하여 그래프 합성곱 연산을 수행하여 가장 최적의 다음 레이어를 설계한다.

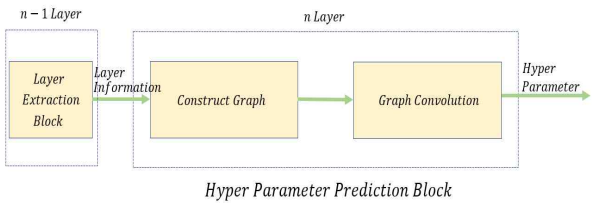


그림 5. 하이퍼 파라미터 추정 그래프 합성곱 신경망 블록 구조
Figure 5. Hyperparameter estimation graph convolutional neural network block structure

그림 5에 보인 바와 같이 하이퍼 파라미터 추정 그래프 합성곱 신경망 블록은 앞단에서 구성된 그래프의 노드의 정보를 기반으로하여 앞단의 레이어에서 다음 단의 레이어로의 특징 추출 성능을 극대화 할 수 있게 현재 구성된 레이어에서 가장 최적의 파라미터를 생성한다. 그래프는 크게 모델의 전체 레이어를 구성하는 레이어 구성 그래프, 모델의 하이퍼 파라미터를 결정하는 파라미터 그래프로 나뉜다.

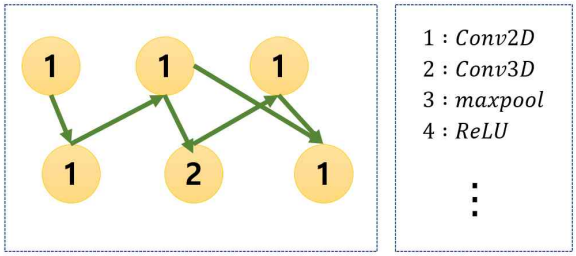


그림 6. 레이어 구성 그래프 예시
Figure 6. Layer composition graph example

그림 6에 보인 바와 같이 레이어 구성 그래프는 하이퍼 파라미터 없이 단순히 레이어의 종류와 연결만을 나타내는 그래프로 연산의 효율성을 위해 레이어의 종류를 숫자로 맵핑하여 기록한다. 이러한 그래프를 통해 다음 레이어 도출 그래프 합성곱 신경망 블록은 모델을 직관적으로 보고 레이어를 결정한다.

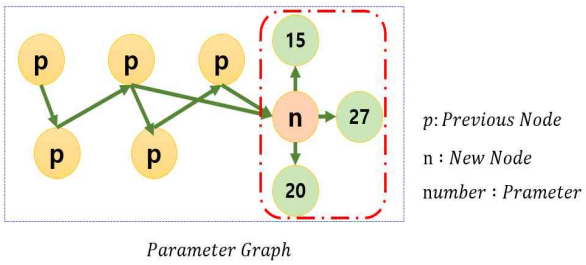


그림 7. 파라미터 그래프 예시
Figure 7. Parametric graph example

그림 7에 보인 바와 같이 파라미터 그래프는 전 단계에서 구성된 레이어와 파라미터 그래프를 기반으로 하여 그래프를 생성한다. 위에서 구성된 레이어 구성 그래프는 단순 연결 정보와 레이어 정보만 담고 있어서 정보량이 매우 적어 하이퍼 파라미터 추정 그래프 합성곱 신경망 블록이 제대로 추론을 하기 어렵다. 따라서 파라미터 그래프는 단순 레이어 정보와 이전 레이어의 노드에서 파라미터 정보를 가져와 새로운 그래프를 생성한다.

그림 8에 보인 바와 같이 이때 가져온 파라미터 정보는 이전 레이어 노드를 주변으로 1차 연결로 구성된다. 위와 같은 그래프 합성곱 연산을 반복하면 n-m-k 레이어의 그래프 노드 k의 정보가 n-m 레이어의 m 노드에

반영되게 된다. 이러한 특성을 이용하여 연산량 및 모델의 크기를 감소시키기 위해 파라미터 그래프는 m번째 연결까지만 유지하고 이전 연결을 삭제하는 구조를 가진다.

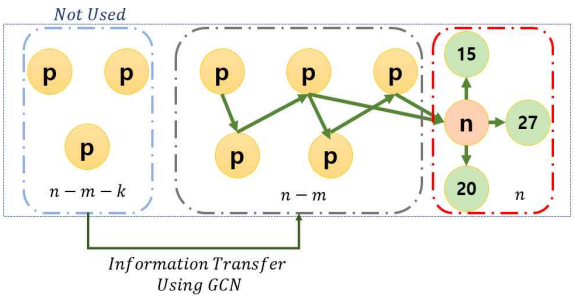


그림 8. 파라미터 그래프의 연결 삭제 예시
Figure 8. Example of deleting connection in parameter graph

## V. 결 론

본 논문은 순환 신경망 계열의 모델 구조가 태생적으로 가지는 고차원 시간 축 문제와 공간적 탐색의 범위 문제를 해결하는 방안을 제시한다. 레이어 도출 그래프 합성곱 신경망 블록은 레이어의 종류 및 연결 정보만을 포함하는 레이어 구성 그래프를 사용하여 모델에게 간략한 정보만 전달하여 직관적인 순차적 레이어를 생성한다. 하이퍼 파라미터 추정 그래프 합성곱 신경망 블록은 레이어의 구조 및 상세 정보를 시공간 축으로 압축하여 생성한 파라미터 그래프를 이용하여 이전 구성된 레이어의 요약된 정보와 이전 레이어의 상세 정보를 반영하여 최적의 파라미터를 도출한다.

본 논문에서 우리는 위와 같은 두 가지 그래프 합성곱 신경망 블록 및 그래프 생성 알고리즘을 통하여 기존 신경 구조 탐색이 가지는 구조적 한계를 해결하는 방안을 제안하였다. 본 논문이 제안하는 GC-NAS를 통하여 신경망 구조 탐색에 그래프 합성곱 신경망을 적용하는 연구가 활발히 이루어질 수 있는 계기가 될 수 있기를 기대한다.

## References

[1] Barret Zoph, and Quoc V. Le. “Neural architecture search with reinforcement learning.” arXiv preprint arXiv:1611.01578, 2016. DOI: https://doi.org/10.48550/arXiv.1611.01578

[2] J. J. Hopfield. “Neural networks and physical systems with emergent collective computational abilities.” Proceedings of the national academy of sciences, 79.8: 2554-2558, 1982. DOI: https://doi.org/10.1073/pnas.79.8.2554

[3] Kunihiko Fukushima and Sei Miyake, “Neocognitron: A self-organizing neural network model for a mechanism of visual pattern recognition,” Competition and cooperation in neural nets. Springer, Berlin, Heidelberg, p. 267-285, 1982. DOI: https://doi.org/10.1007/978-3-642-46466-9_18

[4] Se-Jin Choi, and Jun-Mo Jung. “A Method for accelerating training of Convolutional Neural Networks,” The Journal of the Convergence on Cultural Technology, Vol. 3, No. 4, pp. 171-175, 2017. DOI: 10.17801/JCCT.2017.3.4.171

[5] KIPF, Thomas N. and WELLING. Max. “Semi-supervised classification with graph convolutional networks.” arXiv preprint arXiv:1609.02907, 2016. DOI: https://doi.org/10.48550/arXiv.1609.02907

[6] Rana, Amrita. and Kyung-Ki KIM. “Robust architecture search using network adaptation.” Journal of Sensor Science and Technology, 30.5: 290-294, 2021. DOI: https://doi.org/10.46670/JSST.2021.30.5.290

[7] BOZINOVSKI, and Stevo. “Reminder of the first paper on transfer learning in neural networks,” Informatica, Vol 44. No 3, 2020. DOI: https://doi.org/10.31449/inf.v44i3.2828

[8] DENG. Jia. Dong Wei, and Socher Richard “Imagenet: A large-scale hierarchical image database,” 2009 IEEE conference on computer vision and pattern recognition. Ieee, p. 248-255, 2009. DOI: 10.1109/CVPR.2009.5206848

[9] G. Hinton, O. Vinyals, and J. Dean. “Distilling the knowledge in a neural network.” arXiv preprint arXiv:1503.02531, Feb 2015.

[10]Mingxing Tan, and Quoc Le, “Efficientnet: Rethinking model scaling for convolutional neural networks,” International conference on machine learning. PMLR, p. 6105-6114, 2019.

[11]A. Mirhoseini, A. Goldie, “A graph placement methodology for fast chip design.” Nature, 594.7862: 207-212, 2021.

[12]S. Yan, Y. Xiong, and D. Lin, “Spatial temporal graph convolutional networks for skeleton-based action recognition,” Thirty-second AAAI conference on artificial intelligence, 2018.

[13]W. Peng, X. Hong, and H. Chen, “Learning graph convolutional network for skeleton-based human action recognition by neural searching,”

Proceedings of the AAAI Conference on Artificial Intelligence, p. 2669-2676, 2020. DOI: https://doi.org/10.1609/aaai.v34i03.5652