전파 특성에 기반한 제로샷 학습을 위한 신경망 차원 적응 기법

서연주, 조혜진, 김정환, 이웅희

동국대학교

seo_yeonju@dgu.ac.kr, 2022111851@dgu.ac.kr, kjwh98@dgu.ac.kr, woongheelee@dongguk.edu

Wave Propagation-Inspired Neural Network Dimension Adaptation Techniques: A Zero-Shot Learning Perspective

Yeon-Ju Seo, Hye-Jin Cho, Jung-Hwan Kim, Woong-Hee Lee

Dongguk Univ.

요 약

본 논문은 학습 데이터 셋 구축의 한계 및 기존 학습 데이터와 차원이 다른 데이터가 입력됨에 따른 학습 비용 증가에 주목하여 제로 샷 러닝 기반의 새로운 신경망 적응 기법을 제안한다. 사전 학습된 신경망을 추가 학습 없이 더 높은 차원에서도 쓸 수 있게 geometric sequence decomposition와 sinc 함수 기반 보간 두 가지 기법을 이용해 파라미터(가중치/바이어스)를 확장하는 방식으로 신경망의 입력 차원을 확장한다. 복원과 잡음 제거, 두 모델을 통한 시뮬레이션으로 제안 기법을 통해 확장한 신경망과 재학습을 수행한 신경망을 비교하여 신경망의 입력 차원을 확장할 때 추가적인 학습 없이도 안정적인 성능을 보임을 확인한다. 이로써 사전 학습된 신경망의 재사용 측면에서의 본 논문이 제안하는 기법이 가지는 강점이 기대된다.

Ⅰ. 서 론

인공지능 분야에서, zero-shot learning (ZSL)이란 한 번도 본 적 없는 데이터를 분류 가능하도록 학습하는 것이다. 이러한 학습 방법은 학습에 사용되는 데이터가 없어도 유용한 패턴이나 결과를 도출하기 때문에 학습 데이터가 제한적이거나 레이블 확보가 어려운 환경에서 특히 유용하다 [1]. 반면 few-shot 혹은 many-shot learning 등은 제공되는 학습 데이터의 양에 따라 구분되며, 일반적으로 데이터가 많을수록 성능이 향상된다. 그러나 대규모 데이터 수집 및 라벨링은 높은 시간·비용을 요구하고, 개인정보 보호 및 보안 등의 이유로 데이터 접근 자체가 제한되는 상황도 빈번하게 발생한다 [2].

또한 학습시킨 신경망을 실제 환경에 적용할 때, 데이터 수집 방식이나 환경적인 제약 조건이 달라짐에 따라 기존 학습 데이터의 차원과 다른 새로운 데이터가 들어오는 경우도 발생할 수 있다. 일반적인 신경망은 입력 차원이 고정되어 있어, 신경망으로 들어오는 데이터의 차원이 변하면 이에 맞추기 위해 재학습 과정을 거쳐야 한다. 즉 기존에 학습시킨 신경망을 재사용하지 못하고 새롭게 신경망을 구성해야 한다. 이처럼 기존 신경망의 입력 차원과 다른 데이터가 들어오는 경우, 그 과정에서 또 다시 학습 데이터 확보와 같은 학습 비용이 요구된다는 근본적 문제가 존재한다.

방대한 양의 학습 데이터 확보의 어려움, 데이터 차원 변화에 따른 학습 비용 증가라는 두 가지 문제는 추가 학습 과정 없이도 기존에 학습된 신경망을 재사용하여 입력 차원 변화에 대응할 수 있는 새로운 기법의 필요성을 부각시킨다. 따라서 본 논문에서는 사전 학습된 기존 신경망의 파라미터를 재구성하여, 추가 학습 없이 해당 차원에 맞게 변환함으로써 새로운 입력 차원을 처리할 수 있는 파라미터 확장 기법을 제안한다. 본 논문에서 가장 핵심이 되는, 파라미터를 재구성하는 확장 기법은 geometric sequence decomposition (GSD) [3]과 sinc 함수 기반 보간, 두 가지 방법을 적용하여 파라미터를 확장한다.

Ⅱ. 본론

가. 등비수열 분해 기반 보간 기법

GSD를 기반으로 길이 L인 인코더/디코더 가중치 벡터를 길이 cL로 확장한다. 길이 L의 가중치 벡터 $\mathbf{w}\in\mathbb{C}^{L}$를 K개의 geometric sequence의 합으로 모델링한다.

$$\mathbf{w}[n] = \sum_{i=1}^{K} a_i r_i^n,\ \forall n = \{0,1,\dots,L-1\}. \tag{1}$$

[3]에 따라 K개의 중첩된 관측 시퀀스의 초항 $\{a_i\}_{i=1}^{K}$과 공비 $\{r_i\}_{i=1}^{K}$를 추정하기 위해서는 $2K$개의 값이 필요하다. 이에 본 논문에서는 길이 L인 가중치 벡터 $\mathbf{w}$가 $L \geq 2K$을 만족하도록 $K=L/2$으로 설정하였으며, 이에 기반하여 초항과 공비를 추정하는 과정을 수행한다.

먼저 K개의 중첩된 관측 시퀀스로부터 $(K+1)$개의 연속 꼭짓점 $\mathbf{v}_0,\dots,\mathbf{v}_K$을 다음과 같이 구성한다.

$$\mathbf{v}_i = [\mathbf{w}(i),\dots,\mathbf{w}(i+K-1)]^T,\ \forall i = \{0,1,\dots,K\}. \tag{2}$$

이 $(K+1)$개의 꼭짓점들 중에서 K개를 lexicographical 조합으로 뽑아 길이 $(K+1)$인 K-simplexes의 집합 $\aleph$을 만든다. 그리고 simplex의 부피 $\Lambda(\aleph[p])$를 계산하여 이를 계수로 갖는 K차 다항식을 만들면, 해당 다항식의 근이 공비 $\mathbf{r}$이 된다.

$$\sum_{p=0}^{K} \Lambda(\aleph[p])(-r)^{K-p} = 0. \tag{3}$$

$\mathbf{r}$을 기반으로 만든 행렬 $\mathbf{R}\in\mathbb{C}^{K\times L}$을 다음과 같이 구성하고, $\mathbf{a}=\mathbf{R}^{+}\mathbf{w}$를 풀어 초항 $\mathbf{a}$를 추정한다.

$$\mathbf{R}[i,j] = (\mathbf{r}[i])^{j},\quad \forall i,j\in\{0,1,2,\dots\}. \tag{4}$$

추정된 $\mathbf{r}$과 $\mathbf{a}$를 이용해 cL길이의 확장 가중치 벡터 $\mathbf{w}_{\mathrm{GSD}}$를 생성해야 한다. 식 (1)에서 원래 길이 L에서의 한 번 증가하는 단위 r_i를 $r_i' = r_i^{(1/c)}$으로 확장 인자 c에 맞게 변환한다. 그리고 0부터 cL-1까지 인덱스를 확장하여 $\mathbf{w}$를 재구성한다.

$$\mathrm{w}_{\mathrm{GSD}}[m] = \sum_{i=1}^{K} a_i (r_i')^m, \ \forall m = \{0, ..., cL-1\}. \qquad (5)$$

식 (5)와 같이 확장 후, 기존의 w와 에너지가 동일하도록 정규화 과정을 거쳐 에너지를 보존한다. 최종적으로 $\Re\{\mathrm{w}_{\mathrm{GSD}}\}$을 수행하여 가중치 벡터 확장을 마무리한다. w의 확장은 신경망 노드의 수만큼 진행하고 바이어스는 w 확장과 같은 기법으로 한 번만 수행한다.

나. sinc 함수 기반 보간 기법

가중치를 확장하는 두 번째 방법으로 sinc 함수 기반 보간 기법을 사용하여 w를 좀 더 촘촘한 간격으로 확장된 값으로 바꾼다. 이를 위해 보간 행렬 $\mathrm{S} \in \mathbb{R}^{cL \times L}$을 다음과 같이 구성한다.

$$\mathrm{S}(m, i) = \mathrm{sinc}\left(\frac{m-i}{c}\right), \forall m \in \{0, ..., cL-1\}, i \in \{0, ..., L-1\}. \qquad (6)$$

이는 길이 L벡터를 길이 cL로 확장하는 역할을 한다. 인코더 가중치 벡터 w_{en}는 입력 방향으로 확장되어야 하므로 $\mathrm{w}_{\mathrm{en}}^{\mathrm{sinc}} = \mathrm{w}_{\mathrm{en}} \mathrm{S}^T$로 확장하고, 디코더 가중치 벡터 w_{de}와 출력 바이어스 벡터 b_{de}는 출력 차원 방향으로 확장되어야 하므로 각각 $\mathrm{w}_{\mathrm{de}}^{\mathrm{sinc}} = \mathrm{S}\mathrm{w}_{\mathrm{de}}$, $\mathrm{b}_{\mathrm{de}}^{\mathrm{sinc}} = \mathrm{S}\mathrm{b}_{\mathrm{de}}$으로 계산한다. 마지막으로, 확장된 파라미터를 입력 차원이 cL인 신경망의 파라미터로 구성한다. 이로써 L-차원 신경망에서 학습된 파라미터를 두 가지 방법으로 확장하여 추가 학습 없이 ZSL을 기반으로한 확장 신경망을 구성하였다.

다. 제안 기법의 프레임워크

기본 신경망 입력 차원 L	64	신경망 width	32
확장 인자 c	2	신경망 depth	1
확장된 신경망 입력 차원 cL	128	손실 함수	MSE
기본 신경망 학습 데이터 수 N_{train}	100	Optimizer	Adam
테스트 수 N_{test}	5000	학습률	0.001

표 1. 기본 시뮬레이션 설정.

명확한 비교를 위해 본 논문에서 제안하는 기법으로 파라미터를 2배 확장한 zero-shot 모델 외에도, 파라미터를 무작위로 2배 확장하여 구성한 랜덤 zero-shot 모델을 추가해 성능의 하한선을 확인한다. 또한 확장된 입력 차원에 맞도록 신경망을 재구성한 뒤, 기존 신경망이 학습한 데이터 대비 2배 확장된 데이터로 재학습시킨 many-shot 모델을 구성한다. 시뮬레이션으로 many-shot 모델이 학습하는 데이터 수 $N_{\mathrm{train,exp}}$을 변화시키며 저차원(64)대비 2배 확장된 환경에서의 복원 및 잡음 제거 성능을 mean squared error (MSE)로 평가한다. 이로써 제안하는 두 zero-shot 확장 모델이 추가 학습을 수행한 many-shot 모델 대비 어느 학습 데이터 개수 구간에서 우수한지 확인한다.

라.시뮬레이션 결과

그림1, 그림2에서 공통적으로, many-shot 모델은 $N_{\mathrm{train,exp}}$이 증가함에 따라 MSE가 감소하는 추세를 보인다. 반면, 제안하는 GSD/sinc 기반 zero-shot 확장 모델은 추가적인 학습 과정이 요구되지 않아 $N_{\mathrm{train,exp}}$에 상관없이 동일한 MSE 값을 가짐을 확인할 수 있다. 또한 랜덤 모델은 전 구간에서 가장 높은 MSE를 보였으며, 이는 무작위적인 차원 확장이 아닌 GSD/sinc와 같은 구조적 확장이 성능 확보에 핵심임을 뒷받침한다.

그림 1은 $N_{\mathrm{train,exp}}$변화에 따른 오토인코더 복원 성능(MSE)을 비교한 결과이다. GSD 기반 zero-shot 모델은 $N_{\mathrm{train,exp}}$이 100 이하인 구간에서 many-shot 모델 대비 더 낮은 MSE를 보여 소량 학습 환경에서의 우수성을 확인하였다. 반면 sinc 기반 zero-shot 모델은 그림 상단의 가우시안 신호 모델에서는 $N_{\mathrm{train,exp}}$이 50일 때까지는 경쟁력 있는 성능을 보였으나, K-Sparse 벡터에서는 전반적으로 many-shot 모델 대비 열세를 보여 sinc 기반 보간 방식의 한계를 확인하였다.

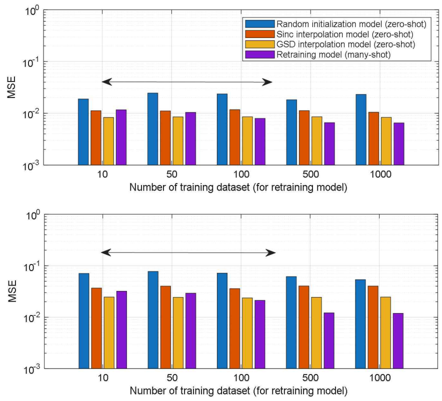

그림 1. 복원 성능 (상단: 가우시안 신호, 하단: K-sparse 벡터, sparsity=3).

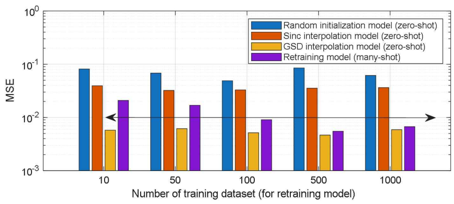

그림 2. K-Sparse 벡터(sparsity=3)의 잡음 제거 성능.

그림 2는 동일한 설정에서 $N_{\mathrm{train,exp}}$변화에 따른 디노이징 오토인코더(DAE)의 잡음 제거 성능을 비교한 결과이다. K-Sparse 벡터에 signal-to-noise ratio (SNR)은 0dB로 가산 백색 가우시안 잡음을 추가하여 노이즈 신호를 모델링하였다. many-shot 모델은 확장된 차원의 데이터를 가지고 재학습을 수행하였음에도 불구하고 전 구간에서 GSD기반 zero-shot 모델보다 높은 MSE를 나타내었다. 이렇게 다양한 환경에서 GSD기반 zero-shot 모델이 잘 작동됨을 확인하였다.

Ⅲ. 결론

본 논문에서는 더 높은 샘플링 레이트나 데이터 자체의 길이가 늘어나는 등 신경망의 입력 차원이 증가하는 경우, 재학습 없이 파라미터를 확장시켜 사전 학습된 기본 신경망을 재사용 하는 기법을 제안하였다. 기본 신경망의 차원보다 작거나 커진, 새로운 차원의 데이터가 들어오더라도 이 때마다 확장 인자 c값을 유동적으로 변화시켜 본 논문에서 제안하는 기법으로 파라미터를 재구성한다면, 새로운 차원의 데이터가 들어올 때마다 해당 차원에 맞는 학습 과정을 수행하지 않아도 되기 때문에 이를 통한 학습 비용감소가 예상된다.

ACKNOWLEDGMENT

본 연구는 2024학년도 동국대학교 신임교원 정착연구비 지원으로 이루어졌음(S-2024-G0001-00025).

참 고 문 헌

[1] F. Pourpanah, M. Abdar, Y. Luo, X. Zhou, R. Wang, C. P. Lim, X.-Z. Wang, and Q. M. J. Wu, “A Review of Generalized Zero-Shot Learning Methods”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 45, no. 4, pp. 4051 - 4070, 2023.

[2] M. Kim, J. Kim, and U. Kang, “Synq: Accurate zero-shot quantization by synthesis-aware fine-tuning”, *ICLR The Thirteenth International Conference on Learning Representations*, 2025.

[3] W.-H. Lee, J.-H. Lee, and K. W. Sung, “Geometric sequence decomposition with *k*-simplexes transform”, *IEEE Transactions on Communications*, vol. 69, no. 1, pp. 94 - 107, 2020.