

A Study on Improving Performance of the Deep Neural Network Model for Relational Reasoning

Hyun-Ok Lee[†] · Heui-Seok Lim^{**}

ABSTRACT

So far, the deep learning, a field of artificial intelligence, has achieved remarkable results in solving problems from unstructured data. However, it is difficult to comprehensively judge situations like humans, and did not reach the level of intelligence that deduced their relations and predicted the next situation. Recently, deep neural networks show that artificial intelligence can possess powerful relational reasoning that is core intellectual ability of human being. In this paper, to analyze and observe the performance of Relation Networks (RN) among the neural networks for relational reasoning, two types of RN-based deep neural network models were constructed and compared with the baseline model. One is a visual question answering RN model using Sort-of-CLEVR and the other is a text-based question answering RN model using bAbI task. In order to maximize the performance of the RN-based model, various performance improvement experiments such as hyper parameters tuning have been proposed and performed. The effectiveness of the proposed performance improvement methods has been verified by applying to the visual QA RN model and the text-based QA RN model, and the new domain model using the dialogue-based LL dataset. As a result of the various experiments, it is found that the initial learning rate is a key factor in determining the performance of the model in both types of RN models. We have observed that the optimal initial learning rate setting found by the proposed random search method can improve the performance of the model up to 99.8%.

Keywords : Relation Network, Relational Reasoning, Text-Based Question Answering, Visual Question Answering

관계 추론 심층 신경망 모델의 성능개선 연구

이 현 옥[†] · 임 희 석^{**}

요 약

지금까지 인공지능의 한 분야인 딥러닝 방법은 구조화되지 않은 데이터로부터 문제를 해결하는 놀라운만한 성과를 이루어왔지만, 인간처럼 여러 상황들을 종합적으로 판단, 그것들의 연관성을 추론하고, 그 다음 상황을 예측하는 수준의 지능을 갖는데 도달하지 못하였다. 최근 발표된 복잡한 관계 추론을 수행하는 심층 신경망은 인공지능이 인간의 핵심 지적 능력인 관계 추론을 보유할 수 있다는 것을 증명하였다. 본 논문에서는 관계 추론 심층 신경망 중에서 Relation Networks (RN)의 성능을 분석 및 관찰해 보고자 Sort-of-CLEVR 데이터 셋을 사용한 시각적 질의응답과 bAbI task를 사용한 텍스트 기반 질의응답 두 유형의 RN 기반 심층 신경망 모델을 구축하여 baseline 모델과의 비교를 통한 성능검증을 하였다. 또한 모델의 성능을 극대화하기 위하여 하이퍼 파라미터 튜닝 등 다양각도의 성능개선 실험으로 관계 추론을 위한 RN 기반 심층 신경망 모델의 성능개선 방법을 제안하였다. 제안한 성능개선 방법은 시각적 질의응답 모델과 텍스트 기반 질의응답 모델에 적용하여 그 효과를 검증하였고, 기존의 RN 모델에서 사용해보지 않았던 Dialog-based LL 데이터 셋을 사용하여 새로운 도메인에서의 제안한 성능개선 방법의 효과를 다시 한 번 검증하였다. 실험 결과 두 유형의 RN 모델 모두에서 초기 학습률이 모델의 성능을 결정하는 핵심 요인임을 알 수 있었고, 제안한 random search 방법에 의해 찾은 최적의 초기 학습률 설정이 모델의 성능을 최고 99.8%까지 향상시킬 수 있다는 것을 확인하였다.

키워드 : 관계 네트워크, 관계 추론, 텍스트 기반 질의응답, 시각적 질의응답

1. 서 론

추리소설이나 범죄 드라마에서 주인공들은 모든 단서의 조각조각의 퍼즐을 맞추어 범인을 찾아내는데, 그들의 범인을 찾아내는 과정을 보면 그 능력에 감탄을 감추지 못할 때가 많다. 범인을 찾아내는 과정에서 모든 단서들 간의 연관성(relations)

※ 이 논문은 2018년도 한국정보처리학회 춘계학술발표대회에서 '파라미터 튜닝을 통한 Relation Networks 성능개선'의 제목으로 발표된 논문을 확장한 것임.

† 준 회 원 : 고려대학교 빅데이터 융합학과 석사

** 정 회 원 : 고려대학교 컴퓨터학과 교수

Manuscript Received : July 6, 2018

Accepted : August 8, 2018

* Corresponding Author : Heui-Seok Lim(limhseok@korea.ac.kr)

을 추론하고, 이들 간 하나의 연결고리를 만들어 범인을 유추해 간다. 이러한 어떤 규칙에 따라 생각하는 사고능력을 우리는 추론이라 일컫는다. 추론은 사고의 관계를 하나의 명제와 다른 하나의 명제의 논리적인 관계로 파악해서 생각하는 것을 의미한다. 이러한 관계 추론(relational reasoning)[1, 2] 능력은 인간이 다른 종과 구별되는 가장 강력하고 유연한 지능(intelligence)의 핵심적인 요소로 간주된다. 지금까지 인공지능의 한 분야인 딥러닝(deep learning) 방법은 영상, 음성 인식 뿐만 아니라 자연어처리(natural language processing) 등 구조화 되지 않은 데이터(unstructured data)로부터 문제를 해결하는 놀라울 만한 성과를 이루어 왔지만, 명시적으로 객체간의 관계를 고려하지 않고 이를 수행해 왔다. 그러나 날로 발전해 가고 있는 인공지능 기술은 최근 이러한 인간의 핵심 지적 능력에 도전장을 내밀고 있다.

관계 추론을 수행하는 Relation Networks[1]와 Visual Interaction Network[3] 이 두 심층 신경망(deep neural networks)은 세상을 객체(objects)와 그들의 관계(relations)라는 체계로 분해하고, 신경망이 피상적으로는 매우 달라 보이지만 근본적으로는 공통관계를 갖는 장면들에 대하여 객체와 관계라는 새로운 결합을 일반화할 수 있는 강력한 추론 능력을 보유할 수 있다는 것을 보여주었다. 이는 심층 신경망이 사람처럼 유연한 사고를 할 수 있게 되었다는 것을 시사해준다.

본 논문에서는 사람처럼 관계 추론을 수행하는 심층 신경망 중에서 Relation Networks(RN)의 성능을 분석 및 관찰해 보고자 시각적 질의응답(Visual Question Answering)[4]과 텍스트 기반 질의응답(Text-based Question Answering) 두 유형의 RN 기반 심층 신경망 모델을 구축하여 baseline 모델과의 비교를 통한 성능검증을 하여 보았다. 또한 모델의 성능을 극대화하기 위하여 하이퍼 파라미터(hyper parameters) 튜닝 등 다양각도의 성능개선 실험으로 관계 추론을 위한 RN 기반 심층 신경망 모델의 성능개선 방법을 제안하였다.

2. 관련 연구

2.1 Relation Networks

기존 시각적 질의응답 문제에 있어 심층 학습의 표준적 접근 방법은 Fig. 1A와 같이 질문(question), 이미지(image) 및 답변(answer)으로 구성된 대규모 데이터 셋(dataset)을 수집하고 질문과 이미지에서 답변으로 직접 매핑하는 단일 신경네트워크를 교육하는 것이었다. 그러나 질의가 Fig. 1B와 같다면 결과는 이와 같은 질문에 대해서는 잘 작동하지 않는다. Fig. 1B의 질문은 다양한 소재와 형태간의 상관관계를 종합해서 답을 추적해 가는 과정이 필요하다. 그러면 이러한 질문에 대한 해결책은 무엇인가?

해결책으로 제시할 수 있는 것은 관계 추론으로써 서로 다른 객체들 간의 관계를 이해하는 것을 학습하는 것이다. 관계 추론은 지능의 기본적 특성으로 간주된다. 이에 답마인드(DeepMind)에서는 이러한 관계 추론을 할 수 있는 RN을 개발하였다. RN을 Convolutional Neural Networks(CNN)과 Long

Short-Term Memory(LSTM)과 결합하여 시각적 질의응답 문제에 대해서 실험을 한 결과, 기존 모델은 76.6% 이하의 성능을, 사람은 92.6%를 그리고 RN을 적용한 모델은 95.5% 성능이 나왔다. 이렇게 사람을 뛰어넘은(super-human) 높은 성능을 보여주는 RN은 간단하며, 다른 모델에 쉽게 붙일 수 있고, 유연한 관계 추론에만 중점을 둘 수 있는 강력함을 가지고 있다[1].

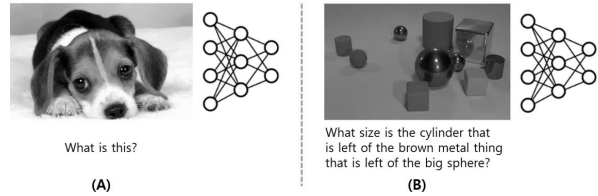


Fig. 1. Visual QA - Standard(A) vs Relational Reasoning(B)

2.2 Relation Networks Architecture

RN의 기본 수식과 시각적 질의응답 및 텍스트 기반 질의응답 RN 모델의 아키텍처(architecture)는 다음과 같다.

1) RN Function

Equation (1)-(2)는 RN의 기본 동작원리를 보여주는 수식이다.

$$RN(O) = f_{\phi} \left(\sum_{i,j} g_{\theta}(o_i, o_j) \right) \quad (1)$$

$$RN(O) = f_{\phi} \left(\sum_{i,j} g_{\theta}(o_i, o_j, q) \right) \quad (2)$$

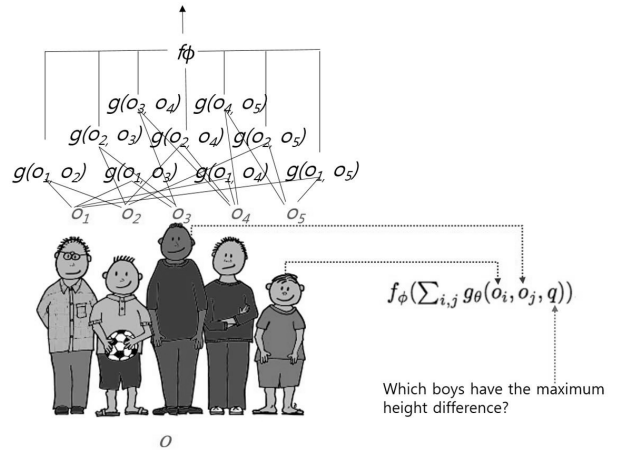


Fig. 2. Visual QA Requiring Relational Reasoning

Equation (1)의 o 는 관계를 알고 싶어 하는 객체(object)를 의미한다. g_{θ} 는 한 쌍의 객체를 입력으로 받아서 관계를 도출하는 함수로 한 쌍의 객체 사이의 관계를 계산하는 다층 퍼셉트론(MLP-Multilayer Perceptron)이다. f_{ϕ} 는 모든 g_{θ} 의 합계를 취하여 모델의 최종 출력을 계산하는 다층 퍼셉트론(MLP)이다.

Fig. 2와 같은 시각적 질의응답 질문의 경우 RN의 수식

은 Equation (2)와 같다. 수식에서 볼 수 있듯이 RN은 하나의 모듈로서 CNN과 LSTM에 플러그인 될 수 있다. CNN로부터 나온 관계를 알고 싶은 각 객체 쌍(objects pair)은 LSTM로부터 나온 질문 벡터(question vector)와 함께 신경망인 g_θ 의 입력으로 사용되어 관계형 네트워크를 학습한다. 이렇게 학습된 g_θ 의 출력이 합쳐져 다른 신경망인 f_ϕ 의 입력으로 사용되고, f_ϕ 는 질문에 대한 답에 대해 최적화된다.

2) Visual QA Architecture

시각적 질의응답 태스크의 RN 모델은 입력으로 들어온 이미지는 CNN에서, 질문은 LSTM에서 처리를 한다. 관계 추론을 위해 입력된 이미지는 CNN으로 처리되어 RN에 대한 객체집합(a set of objects)을 생성되는데, 객체들은 특징 맵 벡터(feature-map vector) 형태로 구성되며 배경, 형태, 질감, 색상 등의 정보를 담고 있다. 질문은 LSTM으로 처리되어 질문 임베딩(question embedding)을 생성한다. 이렇게 생성된 CNN으로 처리한 이미지의 최종 feature map의 객체 쌍과 LSTM으로 처리한 질문의 결과가 다층 퍼셉트론인 g_θ 의 입력으로 사용되고, g_θ 는 각 객체간의 관계를 추론한다. g_θ 가 처리한 모든 조합의 합은 다시 한 번 또 다른 다층 퍼셉트론인 f_ϕ 의 입력으로 들어가게 되고, f_ϕ 는 질문에 대한 최종 답을 내게 된다. 즉 RN은 인코딩된 질문에 따라 모든 객체의

쌍 전반에 걸쳐 관계를 고려하고, 이러한 모든 관계를 통합하여 질문에 대한 최종 답을 낸다고 할 수 있다.

Fig. 3은 두 객체간의 관계를 학습할 수 있는 CNN, LSTM 그리고 RN이 결합된 시각적 질의응답 RN 모델의 아키텍처이다.

3) Text-based QA Architecture

시각적 질의응답 태스크의 RN 모델에서 이미지가 CNN으로 처리되어 객체집합으로 생성된 것과 달리 bAbI 스위트(suite)와 같은 자연어 기반의 질의응답 태스크의 RN 모델은 LSTM으로 처리한 문장의 최종 상태가 객체집합으로 변환된다. 즉 시각적 질의응답 태스크의 RN 모델에서는 CNN feature map이, 텍스트 기반의 질의응답 태스크의 RN 모델에서는 LSTM feature map이 관계를 알고 싶은 객체가 된다. 그러나 질문은 시각적 질의응답 태스크 RN 모델의 질문 인코딩 과정과 마찬가지로 또 다른 LSTM으로 처리되어 질문 임베딩을 생성한다. 텍스트 기반 질의응답 태스크 RN 모델에서는 LSTM으로 처리한 문장의 최종 feature map의 객체 쌍과 또 다른 LSTM으로 처리한 질문의 결과가 다층 퍼셉트론인 g_θ 의 입력으로 사용된다. 마지막으로 g_θ 가 처리한 모든 조합의 합은 또 다른 다층 퍼셉트론인 f_ϕ 의 입력으로 들어가게 되고, f_ϕ 는 질문에 대한 최종 답을 내게 된다. Fig. 4는 자연어로 된 문장 처리를 위한 LSTM, 질문에 대한 처리를 하는 LSTM 그리고

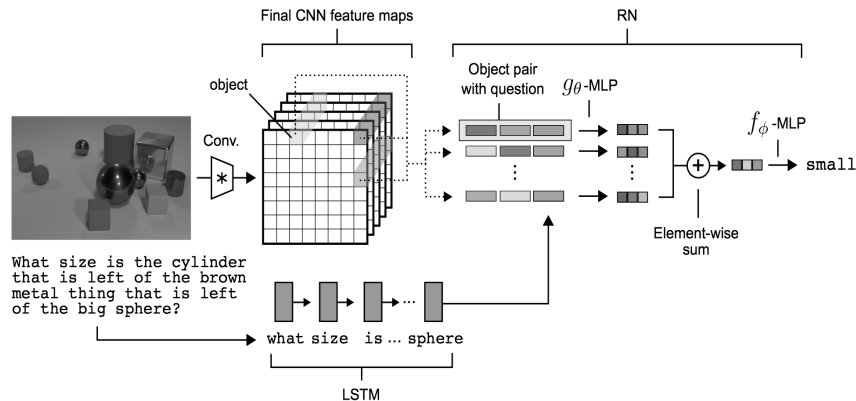


Fig. 3. Visual QA Architecture with RN (Adam Santoro et al., 2017, Figure 2)

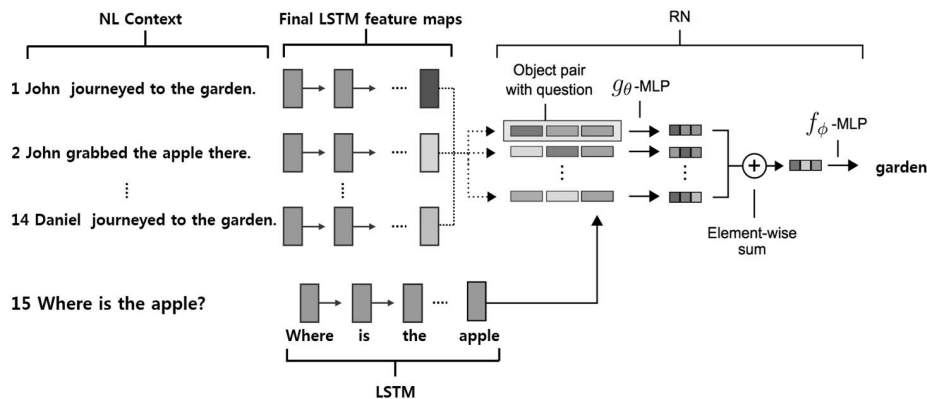


Fig. 4. Text-based QA Architecture with RN

관계를 조합한 후 질문에 대한 최종 답을 산출하는 RN이 결합된 텍스트 기반 질의응답 RN 모델의 아키텍처이다.

3. RN 기반 심층 신경망 모델

본 논문에서는 Fig. 5와 같이 RN 기반 심층 신경망 모델을 구축하여 다양한 도메인에서 RN의 성능을 검증하고, 더 나아가 RN 모델의 성능개선 방법을 제안한다.

다양한 도메인에서의 RN의 성능을 검증하고자, 시각적 질의응답 RN 모델과 텍스트 기반 질의응답 RN 모델 두 유형의 모델을 구축한다. 시각적 질의응답의 RN 모델에서는 데이터 셋은 Sort-of-CLEVR[1]을, 성능비교를 위한 baseline은 CNN + MLP 모델을 사용한다. 텍스트 기반 질의응답의 RN 모델에서는 데이터 셋은 bAbI 태스크(bAbI task)[5]을, 성능비교를 위한 baseline은 End-To-End Memory Networks[6] 모델을 사용한다.

두 유형의 RN 모델은 모델의 성능을 극대화하기 위하여 성능개선 핵심 방법 중에서 오버피팅(overfitting)의 대응책인 배치 정규화(batch normalization)[7], 드롭아웃(dropout)[8] 그리고 최적의 매개변수 값을 찾는 매개변수 갱신법(optimization) 중 하나인 Adaptive Moment Estimation(Adam)[9] 기법을 적용하도록 한다.

또한 본 논문에서는 RN 기반 심층 신경망 모델의 성능개선 방법으로 모델 구축 시 사용한 성능개선 핵심 방법 외의 배치 사이즈(batch size), 초기 학습률(learning rate) 등의 하이퍼 파라미터(hyper parameters)[10] 튜닝방법을 제안한다.

제안한 성능개선 방법은 성능검증을 위해 구축한 시각적 질의응답 RN 모델과 텍스트 기반 질의응답 RN 모델에 적용하여 그 효과를 검증한다. 또한 기존의 RN 모델에서 사용해

보지 않았던 Dialog-based LL 데이터 셋(The Dialog-based Language Learning dataset)[11]을 추가 사용하여 새로운 도메인에서의 제안한 성능개선 방법의 효과를 다시 한 번 검증한다.

3.1 성능 개선 방안 I

RN 기반 심층 신경망 모델은 신경망의 성능개선 핵심 방법 중에서 배치 정규화, 드롭아웃 그리고 매개변수 갱신법 중 하나인 Adam 기법을 적용한다.

배치 정규화는 2015년 제안된 방법으로 학습속도를 개선하고, 초기값에 크게 의존하지 않으며, 오버피팅 억제 효과의 장점을 가지고 있다. 배치 정규화의 기본 아이디어는 각 층에서 활성화 값이 적당히 분포되도록 조정하기 때문에 배치 정규화 계층을 각 신경망에 삽입한다. 즉 배치 정규화를 신경망에 적용 시, hidden layer에 들어가기 전에 배치 정규화 계층(batch normalization layer)을 더해줘서 입력 값을 수정해 준 뒤 이렇게 수정된 값을 활성화함수(activation function) 앞 또는 뒤에 넣어주는 방식을 사용한다. Equation (3)-(5)의 정규화 작업이 끝나면 배치 정규화 계층마다 이 정규화 된 데이터에 Equation (6)의 확대(γ)와 이동(β) 변환을 수행한다.

$$\mu_B \leftarrow \frac{1}{m} \sum_{i=1}^m x_i \tag{3}$$

$$\sigma_B^2 \leftarrow \frac{1}{m} \sum_{i=1}^m (x_i - \mu_B)^2 \tag{4}$$

$$\hat{x}_i \leftarrow \frac{x_i - \mu_B}{\sqrt{\sigma_B^2 + \epsilon}} \tag{5}$$

$$y_i \leftarrow \gamma \hat{x}_i + \beta \equiv BN_{\gamma, \beta}(x_i) \tag{6}$$

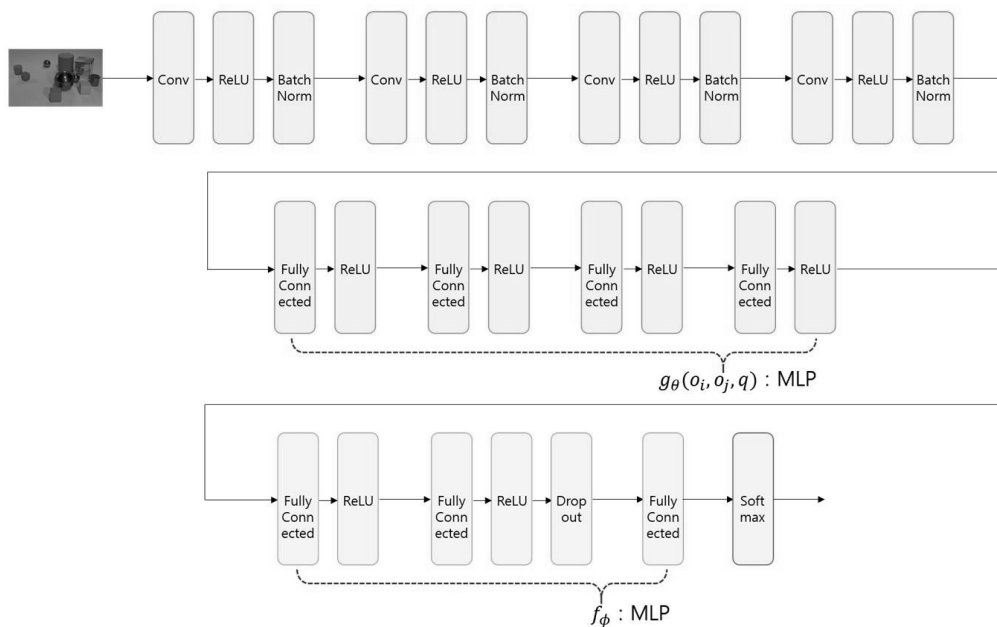


Fig. 5. Our RN-based Visual QA Model Architecture

RN 기반 심층 신경망 모델에서는 Fig. 6에서 볼 수 있듯이 각 층의 활성화함수 ReLU의 활성화 값 분포가 적당히 퍼지면서 학습이 원활하게 이루어지도록 배치 정규화를 사용한다.

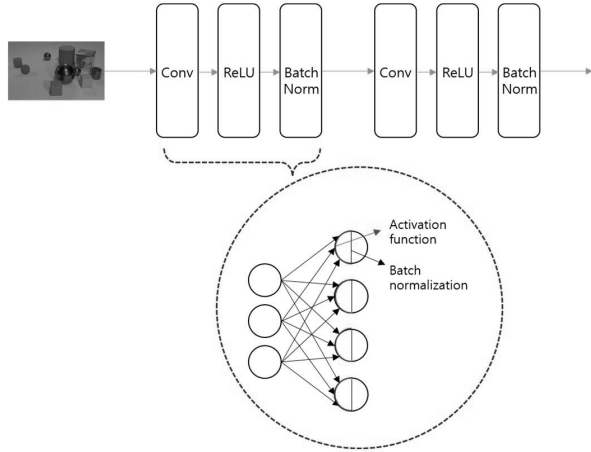


Fig. 6. Neural Network with Batch Normalization

오버피팅을 억제하는 방법으로 신경망 모델이 복잡해지면 가중치 감소만으로 대응하기 어려운 경우 드롭아웃이라는 기법을 사용한다. 드롭아웃이란 신경망 일부를 생략하는 것으로 생략한 신경망은 학습에 영향을 끼치지 않는다. 물론 배치 정규화방법으로 드롭아웃의 효과를 가져 올 수 있지만, 본 논문의 RN 기반 심층 신경망 모델에서는 매개변수 갱신 전에 드롭아웃을 50% 사용하여 좀 더 의미 있는 특징(salient feature)들을 추출하여 자유도(표현력)를 높이고 오버피팅을 억제하도록 한다.

신경망 학습의 목적은 손실 함수의 값을 가능한 한 낮추는 매개변수의 최적값을 찾는 것이고 이러한 문제를 푸는 것을 최적화(optimization)라고 한다. 이러한 최적화에는 여러 가지 방법이 있다. 가장 간단한 방법으로는 손실함수의 그래프에서 가장 낮은 지점을 찾아가도록 손실함수의 기울기를 구해 기울어진 방향으로 매개변수의 값을 갱신해나가는 확률적 경사 하강법(SGD)[12]과 이 방법의 단점을 개선한 Momentum[13], AdaGrad[14], Adam 방법이 있다. 본 논문에서의 RN 기반 심층 신경망 모델은 최적화 방법 중에서 가장 큰 장점을 가지고 있는 Adam 기법을 사용하였다.

2015년에 제안된 Adam 기법은 Equation (7)-(8)과 같이 Momentum과 RMSProp를 융합한 듯한 방법이다. Equation (9)-(10)에서 보듯이 파라미터의 편향보정이 진행되는 특징이 있다. 특히 RMSProp과 마찬가지로 Equation (11)에서 보면 매개변수 갱신 시 $\frac{1}{\sqrt{v_t}}$ 을 곱해 학습률을 조정하는데, 이렇게 하면 가장 많이 움직여서 가장 크게 갱신된 매개변수 원소의 학습률은 낮아지게 하고 가장 적게 갱신된 매개변수 원소의 학습률은 높아지게 하는 효과를 낼 수 있다.

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_t \quad (7)$$

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2) g_t^2 \quad (8)$$

$$\hat{m}_t = \frac{m_t}{1 - \beta_1^t} \quad (9)$$

$$\hat{v}_t = \frac{v_t}{1 - \beta_2^t} \quad (10)$$

$$\theta_{t+1} = \theta_t - \frac{\eta}{\sqrt{v_t} + \epsilon} \hat{m}_t \quad (11)$$

본 논문에서의 RN 기반 심층 신경망 모델은 Adam 기법을 사용하여 하이퍼 파라미터의 편향을 보정, 학습의 갱신강도를 적응적으로 조정한다. 학습이 진행되는 동안 10,000 번째 epoch 마다 학습률 감소(learning rate decay)를 수행하여 학습을 진행하면서 학습률을 점차 줄여가는 방법을 사용한다. 즉 처음에는 크게 학습을 하다가 조금씩 작게 학습하고, 개별 매개변수에 적응적으로 학습을 조정하면서 학습을 진행한다.

3.2 성능 개선 방안 II

본 논문에서는 RN 기반 심층 신경망 모델의 성능개선 방법으로 모델 구축 시 사용한 성능개선 핵심 방법 외의 배치 사이즈, 초기 학습률 등의 하이퍼 파라미터 튜닝방법을 제안한다.

신경망은 입력을 한 번에 하나씩 처리하지 않는다. 처리량을 높이려면 데이터를 일괄 처리하는데, 이를 배치 사이즈라고 한다. 배치 사이즈의 크기가 너무 작으면 일괄 처리 내에 많은 변화가 발생하고 training loss 곡선이 많이 바뀐다. 배치 사이즈의 크기가 너무 크면 메모리가 부족 해 지거나 학습과정이 너무 느리게 진행된다. 따라서 적절한 배치 사이즈가 중요하다고 할 수 있다. 본 논문의 RN 기반 심층 신경망 모델은 성능을 극대화 할 수 있는 모델의 적절한 배치 사이즈를 찾기 위해 baseline 기준 배치 사이즈를 단계적으로 변경하여 적용한다.

학습률과 같은 매개변수를 하이퍼 파라미터라고 한다. 이는 가중치(weight)와 편향(bias)과 같은 신경망의 매개변수와는 성질이 다른 매개변수이다. 신경망의 가중치와 같은 매개변수는 훈련 데이터와 학습 알고리즘에 의해 자동으로 획득되는 반면, 학습률과 같은 하이퍼 파라미터는 여러 시범을 통해 가장 잘 학습을 하는 값을 찾는 과정이 필요하다. 학습률이 너무 작으면 학습시간이 길어져서 손실함수가 거의 갱신되지 않고 학습이 끝나고, 반대로 너무 크면 발산하여 학습이 제대로 이뤄지지 않는다. 따라서 적절한 학습률 조정이 중요하다. 본 논문에서는 학습률의 초기 값 선정이 중요하다는 것에 중점을 두고 초기 학습률 선정이 RN 기반 심층 신경망 모델의 성능에 미치는 영향을 살펴본다. 학습률과 같은 하이퍼 파라미터의 최적화 방법은 과학적 방법보다 다소 수행자의 지혜와 직관에 의존하기 때문에 많은 시행착오를 겪는다. 하이퍼 파라미터의 최적화 방법으로는 Heuristic search, Grid search, Random search, Bayesian optimization[15, 16] 방법이 있다. 본 논문에서는 Heuristic search, Random search 방법에 의거, 경험 및 직관 그리고 선형적 지식을 활용하여 초기 학습률 선정을 제안한다.

4. 실험 및 평가

본 논문에서는 시각적 질의응답 RN 모델과 텍스트 기반 RN 모델 두 유형의 모델을 구축하여 다양한 도메인에서 RN 성능을 비교 평가해 본다. 또한 성능개선 핵심 방법을 적용하여 RN 기반 심층 신경망 모델의 성능개선 효과를 검증한다.

실험환경으로는 OS는 Ubuntu 16.04.4 LTS, Graphics processor는 GeForce GTX 1080Ti, 고성능 GPU 가속 응용 프로그램 개발환경을 위한 CUDA Toolkit 그리고 딥러닝 프레임워크로 tensorflow를 사용한다.

4.1 Visual QA Task

1) Sort-of-CLEVR 생성

Sort-of-CLEVR 데이터 셋은 구골의 답마인드가 RN 아키텍처가 일반적 신경망 아키텍처보다 관계형 추론에 더 적합하다는 가설을 증명하기 위하여 구축한 데이터 셋이다. 이 데이터 셋은 CLEVR[17]의 간소화 버전으로 기존 시각적 질의응답 데이터 셋의 이미지 처리의 복잡성을 줄이고 언어학적 모호성과 편향성을 보완한 데이터 셋이다. 본 논문의 실험에서는 시각적 질의응답 태스크에 사용할 데이터 셋으로 이미지의 사이즈가 128×128×3인 10,000개의 이미지를 포함하고 있는 데이터 셋을 생성하였다. 이중 80%는 훈련에 사용하였고, 나머지 20%는 모델의 성능평가를 위하여 별도로 분리하였다. 각 이미지의 객체는 사각형 또는 원형의 무작위로 선택된 모양을 하고 있다. 또한 각 객체는 서로를 식별하기 위하여 6가지의 색상(빨강, 파랑, 초록, 주황, 노랑, 회색) 중 하나의 색상을 지니고 있다. 질문은 각 이미지 당 20개의 질문-10개의 관계형 질문(relational question)과 비관계형 질문(non-relational question)이 생성되었고 각각의 질문의 길이는 11인

이진 문자열(binary string)로 인코딩 되었다. One-hot 벡터로 인코딩 된 질문에는 참조된 객체의 색상(6 bit), 질문의 유형과 하위유형정보(5 bit)를 담고 있다. Fig. 7에서 본 실험에서 구축한 모델이 생성한 이미지 및 질문과 답을 담고 있는 데이터 셋의 정보를 보다 구체적으로 확인할 수 있다. Fig. 7A에 생성된 질문은 “노란색 객체와 가장 가까운 물체는 무엇인가?”로서, 이는 둘 이상의 객체의 속성에 대한 추론을 요구하고, 객체 사이의 상관관계를 종합하여 답을 낼 수 있는 관계형 질문이다. 반면 Fig. 7B에 생성된 질문은 “녹색의 객체는 원형인가 아니면 사각형인가?”로서, 이는 하나의 객체의 속성에만 국한, 객체 사이의 관계를 고려하지 않고 단일 객체의 속성에 대한 추론을 통해 답을 낼 수 있는 비관계형 질문이다.

2) RN 모델 생성

시각적 질의응답 태스크에 대한 RN 모델은 Fig. 8과 같이 구축하였다.

RN 모델은 각각 24개의 커널(kernel)을 가지는 4개의 합성곱 층(convolutional layers)으로 구성, 각 층마다 활성화 함수로 ReLU를 사용하였고, 각 층에서 활성화 값의 적당한 분포를 조정하기 위하여 배치정규화를 적용하였다. 관계를 도출하는 g_{θ} 함수는 256 유닛(unit)으로 구성된 4개 층의 MLP로 구성하였고 각 층마다 활성화 함수로 ReLU를 사용하였다. 모든 g_{θ} 의 합계를 취하여 모델의 최종 출력을 계산하는 f_{ϕ} 는 첫 번째 및 두 번째 층은 256 유닛으로 구성, 세 번째 층은 10 유닛으로 구성된 3개의 MLP로 구성하였다. 또한 각 층마다 활성화 함수로 ReLU를 사용하였고, 특별히 두 번째 층에서는 드롭아웃을 50% 사용하여 좀 더 의미 있는 특징들을 추출하여 자유도를 높이고 오버피팅을 억제할 수 있도록 하였다. 마지막 출력 층에서는 softmax를 사용하였고, f_{ϕ} 를 통과한 최종적으로 나온 결과를 softmax에 통과시켜 질문에 대한 답을 예측하였다.

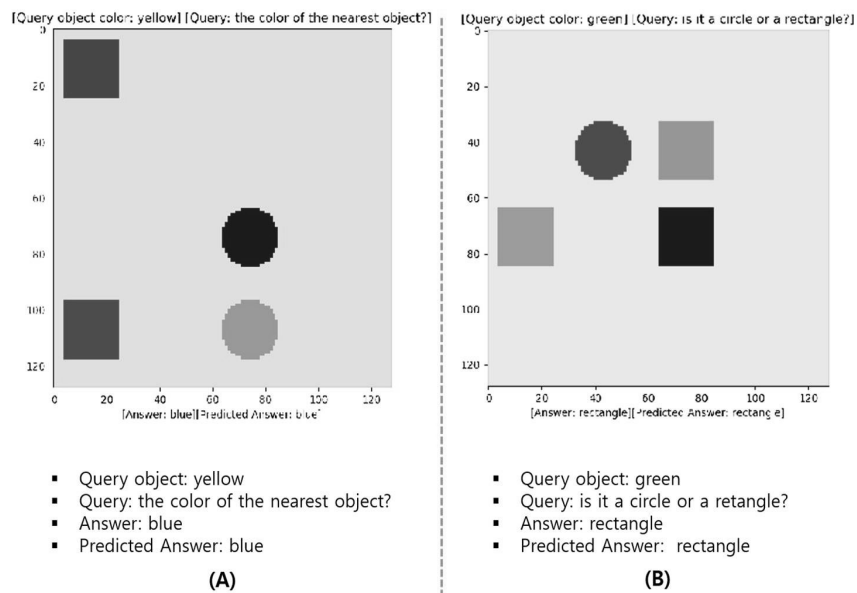


Fig. 7. Relational Question (A) and Non-relational Question (B) Generated on Our Model

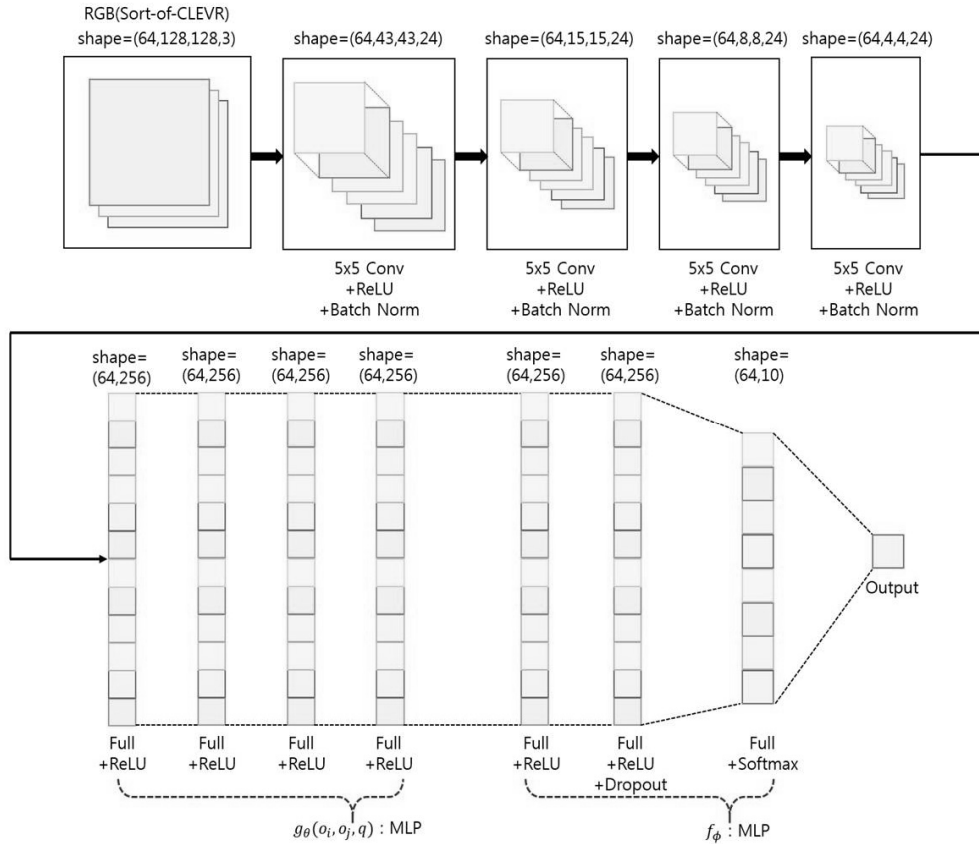


Fig. 8. Our RN-based Model on Visual QA Task

3) Baseline 모델 생성

시각적 질의응답 태스크에 대한 RN 모델의 성능비교를 위한 baseline으로 MLP 기반 모델(CNN + MLP)을 구축하였다. RN 모델과 동일하게 CNN+MLP 모델 또한 각각 24개의 커널을 가지는 4개의 합성곱 층으로 구성, 각 층마다 활성화 함수로는 ReLU를 사용하였고, 각 층에서 활성화 값을 적당히 분포되도록 조정하기 위하여 배치정규화를 적용하였다. 다만, RN 모델에서는 합성곱의 결과를 RN function($g_\theta \rightarrow f_\phi$)에 연결한 것과 달리 CNN+MLP 모델은 합성곱의 결과를 256 유닛으로 구성된 3개 층의 MLP로 연결시켰다. RN 모델과 동일하게 MLP의 2번째 층에서 50%의 드롭아웃을 사용하고 마지막 출력 층에서는 softmax를 사용, 최종적으로 나온 결과를 softmax에 통과시켜 질문에 대한 답을 예측하였다.

4) 성능비교 실험 및 평가

RN 모델 및 baseline으로 사용한 CNN+MLP 모델 모두 학습과정의 배치 사이즈는 64로, epoch은 200,000으로 설정하였다. 학습도중 10 epoch 별로 훈련 데이터와 시험 데이터를 대상으로 정확도(accuracy)를 기록하였다. 모델의 초기 학습률은 $1e^{-4}$ 로 설정하였고, 10,000 step마다 학습률 감소율 0.5로 설정하여 학습을 진행하면서 학습률을 점차 줄여나갔다. RN 모델 및 baseline으로 사용한 CNN+MLP 모델 모두 성능 비교에 사용한 질문은 관계형 질문과 비관계형 질문 두 가지

유형을 사용하였다.

모델의 성능을 극대화하기 위한 방법으로 배치 정규화, 드롭아웃 그리고 Adam optimizer 를 사용하여 구축한 RN 모델은 Table 1에서 볼 수 있듯이 관계형 질문 및 비관계형 질문 모두에 대해서 baseline으로 사용한 CNN+MLP 모델보다 높은 성능을 보였다. 그러나 RN 모델 구축 시 앞에서 나열한 성능개선을 위한 핵심방법을 적용했음에도 RN 모델의 성능이 훈련과정에서의 검증한 수치와 최종 범용 성능 평가에서의 검증한 수치에서 많은 차이를 보였다.

앞에서 적용한 성능개선 방법 외에 RN 모델의 성능개선 방법이 요구됨을 알 수 있다.

Table 1. Performance Comparison Between RN and Baseline on our Visual QA Task

Question type	Model	Performance (%)	
		RN -batch size:64, -learning rate:1e-4	CNN+MLP -batch size:64, -learning rate:1e-4
Non-relational question		61.9	59.5
Relational question		44.5	42.0
Average		55.0	52.5

4.2 하이퍼 파라미터 튜닝을 통한 성능개선 실험

1) 하이퍼 파라미터 튜닝

RN 기반 심층 신경망 모델의 성능개선 방법으로 제안한 하이퍼 파라미터 튜닝을 통해 시각기반 질의응답 RN 모델의 성능개선을 시도해 보았다. 튜닝에 사용한 하이퍼 파라미터로는 배치 사이즈와 초기 학습률을 정하였다.

적절한 배치 사이즈를 찾기 위해 본 논문의 앞서 수행한 시각기반 질의응답 태스크 성능비교 실험에서 구축한 RN 모델을 baseline으로 정하였다. Baseline 모델은 초기 학습률이 $1e^{-4}$ 이고 배치 사이즈가 64이다. baseline 모델 대비 배치 사이즈를 배수로 늘려 배치 사이즈 변경에 따른 성능을 비교 분석해 보았다.

시각기반 질의응답 RN 모델은 성능의 극대화를 위하여 최적화방법 중에서 Adam 기법을 사용, 하이퍼 파라미터의 편향을 보정, 학습의 갱신강도를 적응적으로 조정하였다. 학습이

진행되는 동안 10,000 번째 epoch 마다 학습률 감소를 수행하여 학습을 진행하면서 학습률을 점차 줄여가는 방법을 사용하였다. 그럼에도 모델의 성능이 훈련과정에서의 검증한 수치와 최종 범용 성능 평가에서의 검증한 수치에서 많은 차이를 보였고, RN 모델의 성능이 높지 않게 나왔다. 이에 본 논문에서는 학습률의 초기값 선정이 중요하다는 것에 중점을 두고 초기 학습률 선정이 모델성능에 미치는 영향을 살펴보았다.

본 논문에서는 Heuristic search, Random search 방법에 의거, 경험 및 직관 그리고 선험적 지식을 활용하여 초기 학습률 선정을 위한 실험 및 모델의 성능 검증을 하여 보았다. 적절한 초기 학습률을 찾기 위해 본 논문의 앞서 시각기반 질의응답 태스크 성능비교 실험에서 구현한 RN 모델을 baseline으로 정하였다. Baseline 모델 대비 학습률의 범위를 $1e^{-4} \sim 1e^{-2}$ 로 하여 모델의 성능을 비교 분석해 보았다. 200,000 epoch의 훈련과정이 끝난 후 성능비교 결과는 다음과 같다.

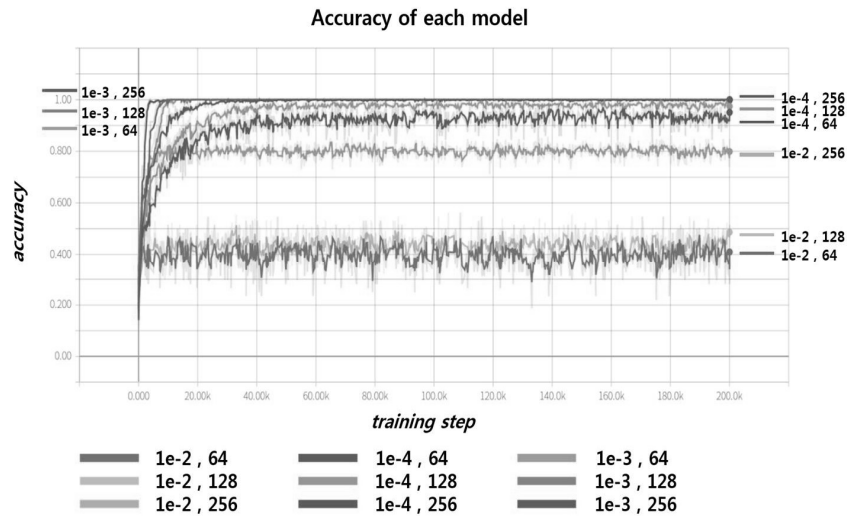


Fig. 9. Accuracy of Each Model with Different Hyper Parameters

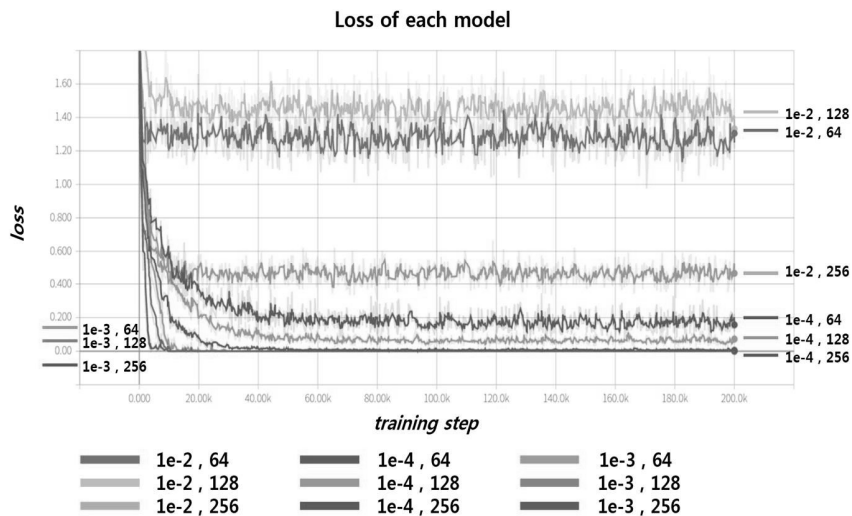


Fig. 10. Loss of Each Model with Different Hyper Parameters

Table 2. Improved Performance by Hyper Parameters Tuning

Learning rate	Batch size	Performance		
		Non-relational question (%)	Relational question (%)	Average (%)
$1e^{-2}$	64	51.7	24.1	40.7
$1e^{-2}$	128	50.5	24.4	40.1
$1e^{-2}$	256	96.3	28.6	69.4
$1e^{-3}$	64	95.4	70.9	85.6
$1e^{-3}$	128	89.2	67.0	80.3
$1e^{-3}$	256	94.6	70.7	85.0
$1e^{-4}$	64	61.6	41.4	53.6
$1e^{-4}$	128	56.8	38.5	49.5
$1e^{-4}$	256	58.8	38.8	50.8

최적화된 모델을 찾기 위한 하이퍼 파라미터 조정에 따른 모델의 성능개선 연구 결과 Fig. 9와 Fig. 10에서 볼 수 있듯이 Heuristic search, Random search 방법에 의해 찾은 초기 학습률이 $1e^{-3}$ 인 모델이 훈련과정에서 성능이 가장 높게 나왔다. 또한 Table 2에서 볼 수 있듯이 초기 학습률이 $1e^{-3}$ 인 모델이 범용실험 성능평가에서도 다른 모델보다 두 배 이상의 성능이 나오는 결과를 보여 주었다.

결론적으로 아무리 훌륭히 설계된 모델이라도 초기 학습률에 의해 모델의 성능이 달라진다는 것과 초기 학습률의 선정이 모델의 성능을 결정하는 중요한 요인임을 알 수 있다.

2) Random Search 방법에 의한 초기 학습률 튜닝

하이퍼 파라미터 튜닝을 통한 성능개선 실험 결과, 다양한 하이퍼 파라미터 중 초기 학습률의 선정이 모델의 성능을 결정하는 중요한 요인임을 알 수 있었다.

학습률의 범위를 $1e^{-6} \sim 1e^{-2}$ 로 정하고 Random search 방법을 사용하여 20,000 epoch의 훈련과정 동안 학습이 잘되는 값의 범위를 관찰해 보았다. 관찰 결과, Fig. 11과 Fig. 12에서 보듯이 학습이 가장 잘 진행될 때의 초기 학습률의 범위는 $3e^{-4} \sim 1e^{-3}$ 이라는 것을 알 수 있다. 최적의 초기 학습률 범위를 이 범위로 축소하고 학습과정을 관찰한다면 최적의 최종 초기 학습률을 찾을 수 있다.

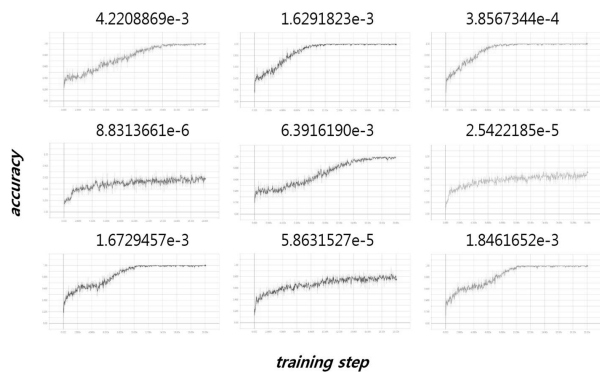


Fig. 11. Accuracy of Each Model with Different Learning Rate Set by Random Search Method

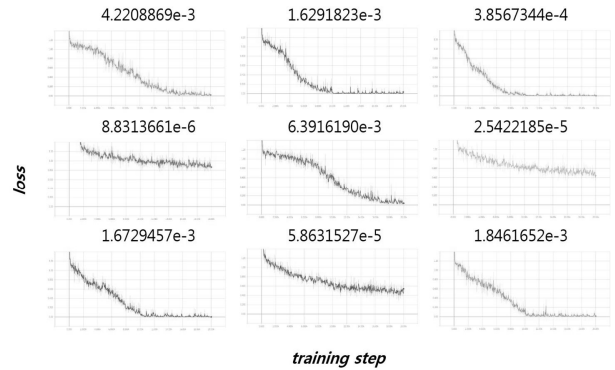


Fig. 12. Loss of Each Model with Different Learning Rate Set by Random Search Method

4.3 Text-based QA Task

1) bAbI 생성 및 모델 구축

bAbI는 순수한 텍스트 기반 QA 데이터 셋이다. bAbI 데이터 셋은 연역적 추론, 귀납적 추론, 계산과 같은 특정 유형의 추론에 해당하는 20여 가지의 태스크로 이루어져 있다. 실험에 사용한 데이터 셋은 bAbI 태스크 중 “Single Supporting Fact”으로써 제공된 924KB의 훈련(training) 데이터와 모델의 성능평가를 위한 96KB의 시험(test) 데이터를 사용하였다.

bAbI 데이터 셋을 사용한 텍스트 기반 질의응답 태스크에 대한 RN 모델은 다음과 같이 구축하였다.

입력으로 들어온 문장은 32unit의 LSTM으로 처리되어 객체를 생성하도록 하였다. 관계를 도출하는 g_{θ} 함수는 64 유닛으로 구성된 4개 층의 MLP로 구성하였고 각 층마다 활성화 함수로 ReLU를 사용하였다. f_{θ} 는 64, 128 유닛으로 구성된 2개의 MLP로 구성하였고 각 층마다 활성화 함수로 ReLU를 사용하였다. 마지막 출력 층에서는 softmax를 사용하였고, f_{θ} 를 통과한 최종적으로 나온 결과를 softmax에 통과시켜 질문에 대한 답을 예측하였다. 앞에서 언급한 텍스트 기반 질의응답 RN 아키텍처에서 볼 수 있듯이 별도의 32unit의 LSTM을 구성하여 질문을 처리하도록 하였다. 또한 드롭아웃을 50% 사용하여 오버피팅을 억제할 수 있도록 하였고 최적화방법으로 Adam optimizer를 사용하였다. 학습과정의 배치 사이즈는 64, 초기 학습률은 $1e^{-4}$, epoch은 200,000으로 설정하였다. 학습도중 100 epoch 별로 훈련 데이터와 시험 데이터를 대상으로 정확도를 기록하였다.

텍스트 기반 질의응답 태스크에 대한 RN 모델의 성능을 비교하기 위한 baseline으로 End-to-End Memory Networks를 구축하였다. Baseline 모델은 문장처리는 LSTM 으로, 드롭아웃은 30%, 최적화방법으로 RMSProp optimizer를 사용하였다.

모델 성능 평가 결과, RN 모델은 정확도가 85.0%, 손실함수는 0.4650의 결과를 보여주었다. 반면 baseline인 End-to-End Memory Networks 모델은 정확도가 99.4%, 손실함수는 0.0481의 결과를 보여 주었다.

결과적으로 baseline 모델보다 RN 모델의 성능이 현저히 떨어져 성능개선이 요구됨을 알 수 있다.

2) 성능개선 실험 및 평가

앞서 수행했던 시각적 질의응답 태스크의 실험에서 초기 학습률에 의해 RN 모델의 성능이 결정되는 것을 확인하였다. 성능개선 방법으로 제안한 초기 학습률 설정이 텍스트 기반 질의응답 bAbI 태스크 RN 모델의 성능개선에도 영향을 주는 지 확인하기 위하여 학습률의 범위를 $1e^{-6} \sim 1e^{-2}$ 로 하여 모델의 성능비교 실험을 수행해 보았다. 최적의 초기 학습률을 찾기 위해 Random search 방법으로 학습이 잘되는 값의 범위를 관찰하고 범위를 좁혀나갔다. 관찰 결과, Table 3에서 보듯이 학습이 가장 잘 진행될 때의 초기 학습률의 범위는 $6e^{-4} \sim 5e^{-4}$ 와 $2e^{-3} \sim 1e^{-3}$ 이라는 것을 알 수 있다.

20,000 epoch의 훈련과정의 모델 범용 성능 평가 결과, 텍스트 기반 질의응답 RN 모델에서도 초기 학습률이 모델의 성능을 결정하는 중요 요인임을 알 수 있다.

Table 3. Comparison of Results on bAbI QA Task Using Different Learning Rate Set by Random Search Method

Learning rate	Performance	
	Accuracy (%)	Loss (%)
2.3025629e⁻³	99.8	0.019473627
1.3385402e⁻³	99.7	0.020856153
1.0048932e⁻³	99.2	0.043623564
6.7598553e⁻⁴	99.3	0.028257005
5.0953512e⁻⁴	98.6	0.045940823
1.2238979e ⁻⁴	53.0	1.0470955
5.4997704e ⁻⁵	45.7	1.6162883
1.7392879e ⁻⁵	42.8	1.8283777
4.7004666e ⁻⁶	33.6	2.3123474
1.5096440e ⁻⁶	24.1	2.2575662

4.4 성능개선 검증 실험

Dialog-based LL 데이터 셋(The Dialog-based Language Learning dataset)을 사용한 모델로 RN 기반 심층 신경망 모델에서 제안한 성능개선 방법을 다시 한 번 검증하는 실험을 해 보았다. 실험에 사용한 데이터 셋은 dialog-based LL 태스크 중 "Imitating an Expert Student"로써 제공된 98KB의 훈련 데이터와 모델의 성능평가를 위한 98KB의 시험 데이터를 사용하였다. 텍스트 기반 질의응답의 Dialog-based LL 태스크에 대한 RN 모델은 bAbI 태스크와 동일한 아키텍처로 구축하였다.

앞서 수행했던 시각적 질의응답 태스크의 실험과 bAbI 데이터 셋을 사용한 텍스트 기반 질의응답 태스크에서 초기 학습률에 의해 모델의 성능이 결정되는 것을 확인하였다. 성능개선 방법으로 제안한 초기 학습률 설정이 Dialog-based LL 태스크에서도 적용되는지 확인하기 위하여 모델의 학습률의 범위를 $1e^{-6} \sim 1e^{-2}$ 로 하여 모델의 성능을 비교 분석해 보았다. 20,000 epoch의 훈련과정이 끝난 후 성능비교 결과는 다음과 같다.

실험결과 학습이 가장 잘 진행될 때의 초기 학습률은 Table 4에서 보듯이 $9e^{-3} \sim 1e^{-3}$ 일 때라는 것을 알 수 있었고, 새로운 테

Table 4. Comparison of Results on Dialog-based LL QA Task Using Different Learning Rate Set by Random Search Method

Learning rate	Performance	
	Accuracy (%)	Loss (%)
9.4494930e⁻³	93.2	0.3993653
6.5708313e⁻³	98.4	0.1106194
1.4750162e⁻³	97.9	0.13057047
1.3382385e ⁻³	80.8	0.69762087
7.9919660e ⁻⁵	46.1	2.0418825
1.8820179e ⁻⁵	46.4	1.2930055
7.7738193e ⁻⁶	38.7	2.0031219
4.9804367e ⁻⁶	36.6	1.6149553
2.1356369e ⁻⁶	23.5	2.2441258

이터 셋인 Dialog-based LL 태스크 RN 모델에서도 초기 학습률이 모델의 성능을 결정하는 중요 요인임을 알 수 있다.

5. 결 론

본 논문에서는 RN 기반 심층 신경망 모델을 구축하여 다양한 도메인에서 RN의 성능을 검증해 보았고, 더 나아가 RN 모델의 성능개선 방법을 시도하여 보았다.

시각적 질의응답의 RN 모델은 Sort-of-CLEVR 데이터 셋을 사용, baseline으로 설정한 CNN + MLP 모델과의 성능 비교를 해 보았다. 텍스트 기반 질의응답의 RN 모델은 데이터 셋으로 bAbI 태스크를 사용, 성능비교를 위한 baseline으로 End-To-End Memory Networks 모델을 사용하였다.

RN 기반 심층 신경망 모델은 성능을 극대화하기 위하여 배치 정규화를 사용하여 각 층의 활성화함수 ReLU의 활성화 값 분포가 적당히 퍼지면서 학습이 원활하게 이루어지도록 하였고, 매개변수 갱신 전에 드롭아웃을 50%를 사용하여 좀더 의미 있는 특징들을 추출하여 자유도를 높이고 오버피팅을 억제할 수 있도록 하였다. 또한 최적화 방법으로 Adam 기법을 사용하여 하이퍼 파라미터의 편향을 보정, 학습의 갱신강도를 적응적으로 조정하였다. 학습이 진행되는 동안 10,000 번째 epoch 마다 학습률 감소를 수행하여 학습을 진행하면서 학습률을 점차 줄여가는 방법을 사용하였다.

이렇게 구축한 RN 모델은 시각적 질의응답 태스크에서 관계형 질문 및 비관계형 질문 모두에 대해서 baseline으로 사용한 CNN+MLP 모델보다 약간 높은 성능을 보여주었다. 반면 텍스트 질의응답 태스크에서는 RN 모델이 baseline으로 사용한 End-to-End Memory Networks 모델보다 현저히 떨어지는 성능을 보여주었다. 또한 RN 모델은 성능개선 핵심방법을 적용했음에도 성능이 훈련과정에서의 검증한 수치와 최종 범용 성능 평가에서의 검증한 수치에서 많은 차이를 보여 적용한 성능개선 핵심 방법 외의 RN 모델의 성능개선 방법이 요구되었다.

RN 기반 심층 신경망 모델의 성능개선 방법으로 모델 구축 시 사용한 성능개선 핵심 방법 외의 배치 사이즈, 초기 학

습률 등의 하이퍼 파라미터 튜닝방법을 제안하였다.

제안한 성능개선 방법은 성능검증을 위해 구축한 시각적 질의응답 RN 모델과 bAbI 데이터 셋을 사용한 텍스트 기반 질의응답 RN 모델에 적용하여 그 효과를 검증하였다. 또한 기존의 RN 모델에서 사용해 보지 않았던 Dialog-based LL 데이터 셋을 추가 사용하여 새로운 도메인에서의 제안한 성능개선 방법의 효과를 다시 한 번 검증하였다.

실험 결과 두 유형의 RN 모델 모두에서 초기 학습률이 모델의 성능을 결정하는 핵심 요인임을 알 수 있었고, RN처럼 아무리 훌륭히 설계된 모델도 초기 학습률에 따라 모델의 성능이 결정된다는 것을 확인하였다.

제안한 Random search 방법에 의해 찾은 최적의 초기 학습률 설정 결과 시각적 질의응답 모델은 최고 99.7%를, 텍스트 기반 질의응답 모델에서는 99.8%까지의 성능 향상을 보여 주었다.

본 논문에서는 Heuristic search, Random search 방법에 의거, 경험 및 직관 그리고 선형적 지식을 활용하여 초기 학습률 선정을 위한 실험 및 그에 따른 모델의 성능을 검증하였다. 추후 연구에서는 학습률의 초기값 선정이 중요하다는 것에 중점을 두고 좀 더 수학적 방법으로 수행하는 Bayesian optimization으로 RN 모델의 초기 학습률의 상세한 조정을 해 볼 것이다.

RN은 기존 신경망에 쉽게 추가할 수 있는 사용이 간단한 play and plug 모듈이다. RN은 이미지나 문장열(series of sentence)과 같은, 구조화 되지 않은 입력을 받아들여 그 안에 포함된 객체간의 관계에 대해 암묵적으로 추론할 수 있다. 객체간의 암묵적 추론을 하는 신경망으로는 RN 외에 물리적인 장면에서의 미래를 예측하는 Visual Interaction Network (VIN)이 있다. RN과 VIN은 관계 추론 접근(relational reasoning approach)에 있어 밝은 전망을 보여주고 있다. 이에 추후 연구에서는 이러한 접근법을 확장, 응용하여 수준 높은 추론모델을 구축하고자 한다. 궁극적으로 이와 같은 연구가 인간의 강력하고 유연한 지능에 대한 핵심적 요소를 이해하는데 많은 도움을 줄 수 있을 것이라고 생각한다.

References

[1] Adam Santoro, David Raposo, David G.T. Barrett, Mateusz Malinowski, Razvan Pascanu, Peter Battaglia, and Timothy Lillicrap, "A simple neural network module for relational reasoning," arXiv: 1706.01427v1, 2017.

[2] David Raposo, Adam Santoro, David Barrett, Razvan Pascanu, Timothy Lillicrap, and Peter Battaglia, "Discovering objects and their relations from entangled scene representations," arXiv:1702.05068, 2017.

[3] Nicholas Watters, Andrea Tacchetti, Théophane Weber, Razvan Pascanu, Peter Battaglia, and Daniel Zoran, "Visual Interaction Networks," arXiv:1706.01433v1, 2017.

[4] Stanislaw Antol, Aishwarya Agrawal, Jiasen Lu, Margaret Mitchell, Dhruv Batra, C Lawrence Zitnick, and Devi Parikh,

"Vqa: Visual question answering," arXiv:1505.00468v7, 2015.

[5] Antoine Bordes, Jason Weston, Sumit Chopra, and Tomas Mikolov, "Towards ai-complete question answering: A set of prerequisite toy tasks," arXiv:1502.05698, 2015.

[6] Sainbayar Sukhbaatar, Arthur Szlam, Jason Weston, and Rob Fergus, "End-To-End Memory Networks," arXiv:1503.08895v5, 2015.

[7] Sergey Ioffe and Christan Szegedy, "Batch Normalization : Accelerating Deep Network Training by Reducing Internal Covariate Shift," arXiv:1502.03167, 2015.

[8] N. Srivastava, G. Hinton. A. Krizhevsky. I. Sutskever, and R. Salakhutdinov, "Dropout : A simple way to prevent neural networks from overfitting," *The Journal of Machine Learning Research*, 15, pp.1929-1958, 2014.

[9] Diederik Kingma and Jimmy Ba, "Adam : A Method for Stochastic Optimization," arXiv: 1412.6980, 2014.

[10] James Bergstra and Yoshua Bengio, "Random Search for Hyper Parameter Optimization," *Journal of Machine Learning Research*, Vol.13, pp.281-305, 2012.

[11] Jason Weston, "Dialog-based Language Learning," arXiv: 1604.06045, 2016.

[12] Sebastian Ruder, "An overview of gradient descent optimization algorithms," arXiv:1609.04747v2, 2016.

[13] Ilya Sutskever, James Martens, George Dahl, and Geoffrey Hinton, "On the importance of initialization and momentum in deep learning," *Proceedings of the 30th International Conference on Machine Learning*, pp.1139-1147, 2013.

[14] John Duchi, Elad Hazan, and Yoram Singer, "Adaptive Subgradient Methods for Online Learning and Stochastic Optimization," *Journal of Machine Learning Research*, Vol.12, pp.2121-2159, 2011.

[15] Jasper Snoek, Hugo Larochelle, and Ryan P. Adams. "Practical Bayesian optimization of machine learning algorithms," arXiv:1206.2944v2, 2012.

[16] Matthias Feurer, Benjamin Letham, and Eytan Bakshy, "Scalable Meta-Learning for Bayesian Optimization," arXiv: 1802.02219, 2018.

[17] Justin Johnson, Bharath Hariharan, Laurens van der Maaten, Li Fei-Fei, C Lawrence Zitnick, and Ross Girshick., "Clevr: A diagnostic dataset for compositional language and elementary visual reasoning," arXiv:1612.06890v1, 2017.



이 현 옥

<https://orcid.org/0000-0003-3540-776X>

e-mail : webpainter@korea.ac.kr

1999년 한국외국어대학교 독일어학(학사)

2018년 고려대학교 빅데이터 융합학과

(석사)

관심분야: 기계학습, 인공지능,

자연어처리, 컴퓨터 비전



임희석

<https://orcid.org/0000-0002-9269-1157>

e-mail : limhseok@korea.ac.kr

1992년 고려대학교 컴퓨터학과(이학사)

1994년 고려대학교 컴퓨터학과(이학석사)

1997년 고려대학교 컴퓨터학과(이학박사)

1997년~현재 고려대학교 컴퓨터학과
교수

관심분야: 자연어처리(NLP), 인공지능, 기계학습, 뇌신경 언어
정보 처리