

기억의 순환 메커니즘을 모델링하기 위한 차원 축소 및 특징값 인코딩 기반 가변형 이미지 표현 방법 연구

강규창*, 조주필^o

Study on Scalable Image Representation through Dimension Reduction and Features Encoding for Memory Circulation Mechanism Modeling

Kyuchang Kang*, Juphil Cho^o

요 약

본 논문은 사람의 기억 순환 메커니즘을 모델링하기 위한 차원축소 및 특징값 인코딩 기반 가변형 이미지 표현 방법에 관한 개념 검증 연구이다. 방법론적으로는 매니폴드 학습 기반으로 입력 데이터의 차원을 축소하고 차원 축소된 변환/잠재공간에서 특징값 기반 인코딩을 통해 이미지를 가변적 비트스트림으로 표현한다. 본 개념 검증용 구현에서 차원축소 및 특징값 추출을 위한 매니폴드 변환 기법으로는 변분 오토인코더를 활용하였고, 차원축소 결과로 추출된 특징값을 인코딩하기 위해서는 사람의 감각기관에서 감각 입력데이터를 정량화하는 개념을 모사한 스칼라 인코딩을 응용하였다. 개념 검증 실험에서는 784차원의 고차원 이미지가 매니폴드 변환으로 10차원으로 축소되고, 축소된 10차원의 부동소수점 특징값을 스칼라 인코딩을 통해 가변 비트의 비트스트림으로 변환이 가능함을 확인하였다. 결과적으로 이미지가 가변 비트의 비트스트림으로 변환됨에 따라 이미지의 인덱싱, 검색, 논리 연산이 가능한 이미지 표현방법으로서 기억의 순환 메커니즘의 기억의 공고화 기능으로 활용할 수 있을 것으로 기대된다.

Key Words : manifold learning, variational autoencoder, scalar encoding, representation learning, image representation

ABSTRACT

This paper is a proof-of-concept study of variable image representation based on dimension reduction and feature value encoding. Methodologically, the dimensionality of the input data is reduced based on the manifold learning, and the image is represented as a variable bitstream through feature-based encoding in the latent space. In the implementation of the concept verification, the variational autoencoder is used as the manifold transformation method for dimension reduction and feature value extraction. We applied a scalar encoding that simulates the concept of quantifying sensory input data in human sensory organs. In the proof-of-concept experiment, it was confirmed that the 784-dimensional image is reduced to 10-dimensional features by manifold transform, and that the reduced 10-dimensional floating-point feature values can be converted into a bitstream of variable bits through scalar encoding. As a result, as the image is converted into a bitstream of variable bits, it is expected that it can be used as a memory consolidation function of the memory circulation mechanism.

* 본 연구는 한국연구재단 (No. NRF-2018R1D1A1B0704923113) 및 군산대학교 대학자체 학술공모과제 연구비 지원으로 수행되었음

[•] First Author : Kunsan National University, School of IT Information and Control Engineering, kc.kang@kunsan.ac.kr, 정희원

^o Corresponding Author : Kunsan National University, School of IT Information and Control Engineering, stefano@kunsan.ac.kr, 정희원
논문번호 : 202003-041-C-RE, Received February 29, 2020; Revised April 30, 2020; Accepted May 14, 2020

1. 서론

인공지능 분야의 새로운 연구 경향으로 판단에 대한 이유를 사람이 이해할 수 있는 방식으로 제시하는 설명 가능한 인공지능(Explainable Artificial Intelligence, XAI)^[1]과 사람의 학습과정을 모사하는 자율성장하는 인공지능^[2]이 새롭게 주목을 받고 있다. 특히 자율성장 인공지능에서는 자가학습이 가능한 기계학습 방법론을 개발하고 있고 또한 인간의 기억 과정을 모사하여 단기 기억과 장기기억의 전환 메커니즘을 소프트웨어로 구현하려고 시도하고 있다. 기억의 순환 메커니즘은 정보를 받아들인 뒤 그것을 저장했다가 나중에 다시 불러오는 일련의 과정으로 정보의 부호화, 저장, 인출 과정으로 나뉜다.

본 논문은 자율성장 인공지능 기술의 한 분야로 기억 순환 메커니즘을 소프트웨어적으로 구현하기 위한 초기 개념 검증연구(PoC, proof of concept)로, 본 연구에서는 특히 사람의 시각 정보에 대한 기억 순환 모델링의 부호화 및 저장 단계의 개념 검증에 초점을 맞추고 있다. 그러나 본 개념 검증연구에서는 사람이라는 요인을 최대한 배제하고 기억의 순환 메커니즘 입력의 하나인 시각 정보를 데이터 관점에서 바라보고 이를 뇌에서 정보의 개념화(conceptualization) 및 공고화(consolidation) 과정을 거쳐 장기기억으로 가져가는 메커니즘으로 활용하기 위해 데이터 처리에 초점을 맞추었다. 따라서 우선적으로 데이터 처리의 관점에서 시각 입력 정보를 차원 축소 및 특징값 인코딩 기반으로 입력 정보를 가변형 이미지로 표현하는 방법을 1차 연구 목표로 정하였다. 본 연구 이후에는 비트 단위로 표현된 가변형 이미지로 정보의 개념화 방법에 관한 개념 검증연구가 진행 예정이고 마지막으로 새로 유입된 입력과 과거 기억 정보를 인출 및 융합하여 기억의 재공고화(reconsolidation) 과정에 관한 개념 검증연구가 진행될 예정이다.

데이터의 관점에서 살펴보면, 일반적으로 데이터는 다양한 속성으로 이루어진 수많은 레코드의 모임으로 간주될 수 있고 각 레코드의 성질 및 특징(feature)은 속성(attribute)의 개념으로 표현될 수 있다. 이 속성의 개수를 차원(dimension)이라고 하면 복잡한 데이터는 차원이 매우 높은 고차원의 데이터가 된다. 데이터의 차원이 높아지면 표현 공간(representation space) 대비 데이터의 밀도가 기하급수적으로 감소하고 중요도가 떨어지는 데이터 특징을 기계학습과 같은 학습 과정에 사용할 경우 과적합(overfitting)에 쉽게 빠질 수 있다. 따라서 차원을 낮추어 데이터 해석을 하기 위한

매니폴드(manifold) 학습 기반의 비선형 차원 축소 기법이 연구되었다.

통상적으로 대용량의 데이터셋을 사용하는 기계학습 문제에서는 특징 공간이 매우 고차원이며 데이터 구조도 매우 복잡하고 다양하다. 따라서 사람이 데이터 구조를 파악하고 변환하는 방법을 설계하는 수작업 기반 방식은 비현실적이며 비지도 학습 알고리즘이 데이터의 구조를 스스로 파악하여 최적의 공간 변환을 알아내는 접근방법을 사용해야 한다. 이때 원래 공간을 다른 공간으로 변환하는 과정을 인코딩, 원래 공간으로 역변환하는 과정을 디코딩이라고 할 수 있다.

정리하면, 데이터의 범위를 줄여서 입력이 이미지라고 가정할 경우, 이미지는 정보의 공백이 많은 고차원의 데이터이다. 따라서 이미지의 주성분 또는 주요 속성 기반으로 차원축소를 할 수 있다. 그러면 기존의 정보를 크게 훼손하지 않으면서 이미지를 얼마만큼 차원 축소 할 수 있을지 살펴볼 필요가 있다. 그리고 차원이 축소된 이미지를 어떻게 표현할 것인가에 대한 방법론도 필요하다.

본 논문은 이러한 요구사항을 데이터 처리 관점에서 풀어보고자 하는 기초 연구로서 차원 축소 방법 및 차원 축소 결과로 만들어진 이미지의 특징값을 인코딩하는 가변형 이미지 표현방법에 관한 개념 검증연구이다. 방법론적으로는 매니폴드 학습 기반으로 입력 데이터의 차원을 축소하고 차원 축소된 변환/잠재공간(latent space)에서 특징값 기반 인코딩을 통해 이미지를 가변적 비트스트림으로 표현한다.

본 개념 검증 구현에서 차원 축소 및 특징값 추출을 위한 매니폴드 변환 기법으로는 변분 오토인코더(VAE, variational autoencoder)^[3]를 활용하였고, 차원 축소 결과로 추출된 특징값을 인코딩하기 위해서는 사람의 감각기관에서 감각 입력데이터를 정량화하는 개념을 모사한 스칼라 인코딩^[4]을 응용하였다.

본 개념 검증 구현의 가용성을 확인하기 위해서 컴퓨터 비전 분야에서 기본적으로 많이 사용되는 MNIST(Modified National Institute of Standards and Technology database)^[5] 데이터셋을 활용하여 제안한 방법론의 가용성을 확인하는 실험을 수행하였다.

본 연구의 선행연구^[6,7]로 매니폴드 학습 기반 특징값을 활용한 카테고리 분류에 대한 사례 연구가 먼저 진행되었다. 이 연구에서는 변분 오토인코더를 활용하여 고차원의 입력 데이터/이미지를 저차원의 특징벡터 공간으로 변환하고 k-최근접 이웃 알고리즘(k-NN)^[8]를 이용하여 카테고리 분류에 대한 실험을 시행하였

다. 이 연구에서 입력 이미지를 차원 축소한 특징값들을 이용하여 분류기 실험을 하였을 때, 원시 데이터를 사용했을 때와 비교하여 분류기 성능이 크게 차이 나지 않는다는 것을 확인하였다. 그러면 왜 이러한 결과가 나왔는지에 대한 직관을 얻기 위해 저차원 변환된 특징값들에 대해 그래프 기반으로 시각화하여 세부적으로 특징값 들의 특성을 살펴보았다. 이 연구에서는 저차원 변환된 특징값들이 변환된 공간에서 어떻게 표현되는지에 대해 분석하였고 이러한 분석을 기반으로 이미지를 다른 형태로 표현하기 위한 표현법에 대한 필요성을 제시하였다.

기억의 순환 메커니즘 관점에서 제안한 가변형 이미지 표현 방법은 기억 순환 과정의 부호화(encoding) 및 저장(storage) 단계에서 기반 기술로 활용될 수 있을 것으로 기대된다.

본 논문은 다음과 같이 구성된다. 2장에서는 제안한 시스템을 설계하기 위한 관련 기술에 대해 살펴보고 시스템 설계에 사용된 기술 내용에 관해 설명한다. 3장에서는 시스템 구현 및 결과에 대해서 세부적으로 설명하고 4장에서는 제안한 방법에 따라 입력 이미지를 비트스트림으로 표현한 예를 설명하고 고찰한다. 마지막으로 5장 결론에서는 개념 검증연구로 진행된 시스템에 대해 정리하고 향후 추가 연구가 필요한 부분에 관해 기술한다.

II. 관련 기술 및 시스템 설계

본 장에서는 먼저 대표적인 인지과학적 기억모델인 에킨슨-쉬프린 모델(Atkinson & Shiffrin model)^[9]에 대해 개념적으로 살펴보고 이를 소프트웨어로 모델링할 때 필요한 기술요소에 관해 설명한다. 그리고 데이터 처리 관점에서 제안한 방법을 설명하기 위해 먼저 입력데이터의 차원축소 방법에 대해 매니폴드 개념에서 출발하여 변분변환 및 변분추론의 개념으로 확장하고 이러한 수학적 변환을 기반으로 하는 변분 오토인코더에 대해서 세부적으로 살펴본다. 다음으로 변분 오토인코더에 의해 입력데이터가 부동 소수점의 저차원 특징값으로 변환되었을 때 이를 정형화하고 저장이나 검색 등의 편의를 위해 비트스트림으로 변환할 인코더 관점에서 스칼라 인코더에 살펴본다.

본 논문의 제안 시스템에서는 이러한 기술들이 연속적으로 연결되어 차원축소 및 특징값 인코딩 기반 가변형 이미지 표현시스템을 제공한다.

2.1 기억 모델

기본적으로 기억은 정보를 받아들인 뒤 그것을 저장했다가 나중에 다시 불러오는 일련의 과정으로 정보의 부호화(encoding), 저장(storage), 인출(retrieval) 과정으로 나뉜다. 부호화 단계는 기억내용을 두뇌에 입력하는 과정이고 저장 단계는 기억내용을 두뇌에 저장하는 과정이며 인출 단계는 기억내용을 두뇌에서 꺼내는 과정으로 볼 수 있다.

그림 1의 에킨슨-쉬프린 모델은 기억 모델 중 고전적인 모델에 속하지만 아직 기억 연구의 많은 부분에 영향을 미치고 있다. 에킨슨-쉬프린 모델에서 ‘단기 기억’과 ‘장기 기억’이라는 개념이 처음으로 등장한다.

기억 순환 관점에서 단기 기억이 장기 기억화 되는 과정을 기억의 부호화 또는 공고화(consolidation)라고 한다. 단기 기억에서 활성화되는 정보들은 새로 입력된 정보만이 아닐 수도 있다. 즉 이전에 이미 장기 기억으로 공고화되었던 기억이 다시 단기 기억의 도마 위에 오르면 (인출), 다시 유연한 상태가 되어 변형과 조작을 할 수 있다. 즉 다시 장기 기억으로의 공고화를 거치게 되는데, 이를 재공고화(reconsolidation)라고 한다.

인지과학적 관점의 기억 순환 모델링을 소프트웨어적으로 모사한다면, 원시 데이터를 입력받고 기억의 단편에 해당하는 특징값을 생성하는 데이터 처리 모듈이 필요하다. 그리고 이를 모아 범주화하는, 즉 소프트웨어적으로 특징값을 클러스터링하고 라벨을 붙여 개념화/범주화하는 모듈이 필요하다. 마지막으로 범주화된 기억의 단편이 인출되고 상호 연합하여 재범주화 되기 위해서는 범주화된 기억의 단편 들이 상호 연합될 수 있는 데이터의 표현방법이 필요하다.

본 논문에서 제안하는 이미지 표현방법은 입력 이미지가 특징값으로 추출되고 클러스터링 되어 논리적 연산이 가능한 비트스트림으로 표현되는 방법으로서 기억 모델의 관점에서는 단기 기억에서 처리되는 데이터 형태라고 할 수 있고 장기 기억으로 저장되는 비트스트림이라고 할 수 있다.

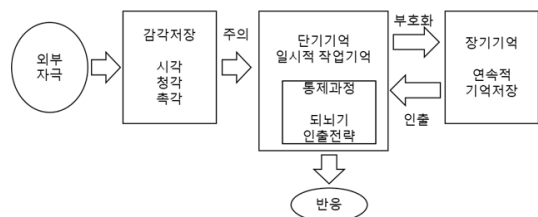


그림 1. 에킨슨-쉬프린의 기억 모델
Fig. 1. Atkinson & Shiffrin memory model

2.2 매니폴드 개념

데이터 분포의 비선형 구조를 직접 고려하는 학습으로 매니폴드 학습이 있다. 고차원 공간에 내재한 저차원 공간을 매니폴드라고 하는데, 수학에서는 매니폴드를 엄격하게 정의하지만, 기계학습에서는 개념적으로 다룬다. 고차원 공간에 주어진 실제 세계의 데이터는 고차원 입력 공간에 내재한 훨씬 저차원 매니폴드 인근에 집중되어 있다^[10].

매니폴드 개념을 그림 2를 통해 설명하면, d 차원 매니폴드는 m 차원 공간에 임베딩되어 있으며 매핑은 식(1)과 같다.

$$f: R^d \rightarrow R^m, \text{ where } d \leq m \quad (1)$$

여기서 주어진 샘플 $x_i \in R^m$ 는 식(2)와 같이 노이즈를 포함하는 수식으로 표현될 수 있다.

$$x_i = f(\tau_i) + \epsilon_i \quad (2)$$

여기서, $f(\cdot)$ 는 임베딩 함수라고 하며, m 은 외적 차원(extrinsic dimension), d 는 잠재공간의 차원 또는 내적 차원(intrinsic dimension)이라고 한다.

주어진 데이터(x_i)에서 $f(\cdot)$ 나 τ_i 를 찾는 것을 매니폴드 학습이라고 한다. 이러한 학습과정에서 우리는 변환을 위해 가정하는 데이터의 확률 분포 $p(\tau)$ 는 매끄럽고(smooth), 균일하게 분포되어 있으며 잡음이 작다는 매니폴드 가설(manifold hypothesis)를 따른다고 전제한다.

데이터 분포의 비선형 구조를 직접 고려하는 매니폴드 학습은 고차원 공간에 내재된 저차원 공간, 즉 변환 결과의 특징값을 결정하기 위해 하이퍼 파라미

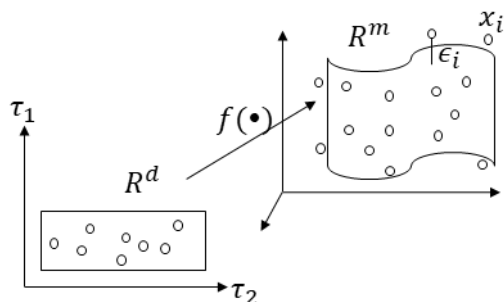


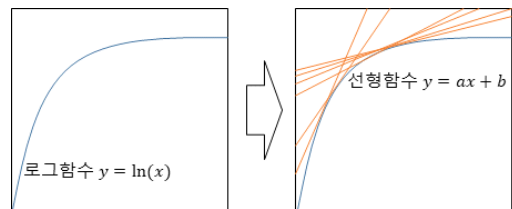
그림 2. 매니폴드 개념 (d 차원 매니폴드는 m 차원 공간에 임베딩되어 있음)
Fig. 2. Concept of manifold (d-dimensional manifold is embedded in m-dimensional space)

터를 찾아가는 과정이다. 이러한 학습 과정에서 매니폴드 변환과 관련된 하이퍼 파라미터가 학습되는데 사용하는 방법론, 예를 들어, VAE(Variational Autoencoder)^[3], AE(Autoencoder)^[11], Isomap^[12], LLE(Locally Linear Embedding)^[13], t-SNE(t-distributed Stochastic Neighbor Embedding)^[14] 등의 방법론을 사용할 수 있고, 각 방법론에 따라 학습해야 하는 하이퍼 파라미터는 각각 달라진다.

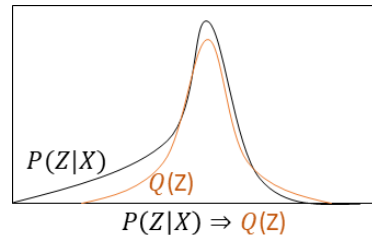
2.3 변분추론과 변분 오토인코더

매니폴드 학습의 이론을 변분추론(variational inference) 관점으로 적용할 경우를 고려할 수 있다.

변분추론은 보통 계산하기 상당히 어려운 사후확률 분포를 다루기 쉬운 확률분포로 근사하는 것을 말한다. 변분추론의 개념은 변분변환(variational transform)을 확률 모델로 확장한 것으로 생각할 수 있다. 변분변환은 예를 들어 그림 3의 (a)에서와 같이 복잡한 비선형 함수인 로그 함수, $y = \ln(x)$ 를 단순한 직선의 선형함수, $y = ax + b$ 로 바꾸어 해석하는 것으로 모든 범위에 대해 만족하는 선형함수를 구하기는 어렵지만 특정한 범위로 한정할 경우 직선의 방정식으로 로그를 표현할 수 있다는 것이다. 변분추론은 이러한 접근법을 확률모델로 확장하여 적용한 것으로 그림 3의 (b)에서와 같이 복잡한 확률밀도함수(probability distribution function), $P(Z|X)$ 를 이미



(a) 변분변환의 개념



(b) 변분추론의 개념

그림 3. 변분변환 및 변분추론의 개념도
Fig. 3. Concept diagram of variational transform and variational inference

잘 알고 있는 간단한 확률 밀도함수, $Q(Z)$ 로 근사화하고 그 파라미터를 학습을 통해 추론할 수 있다는 것이다. 예를 들어 변분추론의 적용에서 실제 데이터의 분포를 이미 잘 알고 있는 정규분포로 근사화하고 학습을 통해 평균과 표준편차를 추론하면 된다.

우리가 다루어야 할 데이터가 이미지라고 가정하면, 이미지는 정보의 공백이 많은 고차원의 데이터로 간주할 수 있다. 따라서 이미지의 주성분 또는 주요 속성 기반으로 차원축소를 할 수 있으므로 이미지는 매니폴드로 표현될 수 있다. 이미지의 경우에는 작은 조각 영상을 다수의 기저 벡터 선형 결합으로 표현하는 희소 코딩 방법론이 있다. 이미지는 폭과 높이를 가지는 2차원 구조인데 이미지 폭의 크기와 같은 행을 높이의 크기만큼 여러 개 쌓아 벡터 구조로 바꿀 수 있다. 예를 들어 28x28 이미지는 28개 컬럼을 가지는 28개의 행을 쌓아 784 차원의 벡터로 만들 수 있다.

자연계에서 발생하는 모든 신호는 특정 함수들의 조합으로 표현할 수 있으며 푸리에 변환에서는 사인 함수와 코사인함수를 사용하여 신호를 표현한다. 이와 같은 맥락으로 이미지에서는 축소된 차원에서 기저벡터의 조합으로 이미지를 표현할 수 있다. 그러나 차원 축소를 위해 적용하는 방법론에 따라 서로 다른 하이퍼 파라미터를 학습해야 한다.

매니폴드는 저차원의 유클리드 공간으로 표현할 수 있으므로 이미지는 평균과 표준편차의 조합으로 표시되는 유클리드 축의 조합으로 표현할 수 있다. 따라서 이미지는 더는 픽셀값의 표현이 아니라 저차원의 유클리드 공간을 생성하는 잠재벡터 (latent vector)로 표현할 수 있는데 킹마(D. P. Kingma)^[3]는 변분추론을 기반으로 하는 변분 오토인코더를 통해 이미지를 잠재벡터로 표현하는 기법 및 학습하는 방법을 제안하였고 그 기본 구조는 그림 4와 같다.

변분 오토인코더를 기반으로 한 매니폴드 변환을 사용할 경우, 학습되는 하이퍼 파라미터인 특징값을 나타내는 잠재변수(z)는 차원수(N), 입력 이미지의 확률분포를 추정하는 가우시안 확률 분포의 평균(μ)과 표준편차(σ), 정규분포에서 샘플링한 임의의 값(ϵ)를 기준으로 하는 식(3)과 같이 계산된다.

$$z_i = \mu_i + \sigma_i \odot \epsilon_i \quad (3)$$

잠재 변수(z)와 변분 오토인코더 구조 관점에서 살펴보면, 인코더는 입력데이터를 차원축소하고 추상화

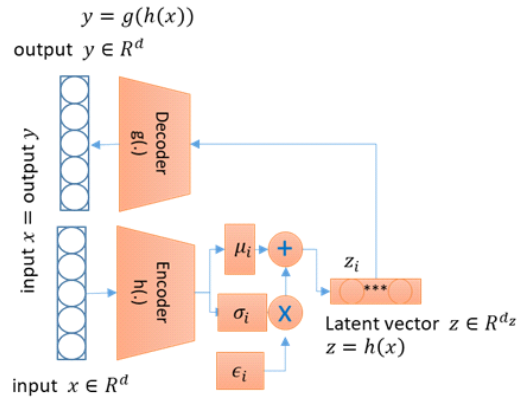


그림 4. 변분 오토인코더의 기본 구조
Fig. 4. Basic architecture of variational autoencoder

하여 잠재적인 특징을 추출하는 역할을 하고, 디코더는 잠재적인 특징을 바탕으로 데이터로 복원하는 역할을 한다. 따라서 학습이 완료된 변분 오토인코더에서는 임의의 잠재변수(z) 값을 디코더에 넣으면 학습에 사용한 입력데이터와 유사한 이미지를 생성할 수 있다.

변분 오토인코더는 입력데이터(x)의 확률분포를 추론하기 위한 인코더 신경망과 입력과 동일한 출력데이터(y , 여기서 입출력을 동일하게 설정하므로 $y = x$)를 생성하는 디코더 신경망으로 구성된다. 잠재변수(z)를 가정하면 인코더 신경망은 관측된 데이터 x 인 이미지를 입력받아 신경망을 통과하고 잠재변수 z 를 만들어낸다. 디코더 신경망은 인코더 신경망이 만든 잠재변수 z 를 활용해 입력데이터 x 를 복원하는 역할을 한다.

2.4 스칼라 인코딩(Scalar Encoding)

뇌에서 시각적 사물을 인식하는 것은 매우 복잡한 구조를 가지며 이미지 자체가 아니라 이미지의 속성이 결합하여 사물을 인식하는 경로와 움직임을 인식하는 경로로 나누어지고 대뇌 신피질에서 인식 정보가 결합된다^[15]. 호킨스(J. Hawkins)는 온인텔리전스^[16]에서 뇌의 신피질에서 시각을 인식하는 것은 최소하게 분포된 뉴런의 발화 패턴으로 인식한다고 주장하고 있다. 그러면 소프트웨어로 대뇌 신피질에서 이미지를 인식하는 방식을 모사하고자 하면 매니폴드 변환에 의한 이미지의 특징값을 인코딩하여 대뇌 신피질에서 뉴런의 발화 패턴으로 표현되는 방식과 유사하게 이미지의 특징값을 표현할 수 있는 방법론에 대해 고려해볼 필요가 있다.

먼저 좀 더 단순한 구조인 청각의 인식메커니즘을

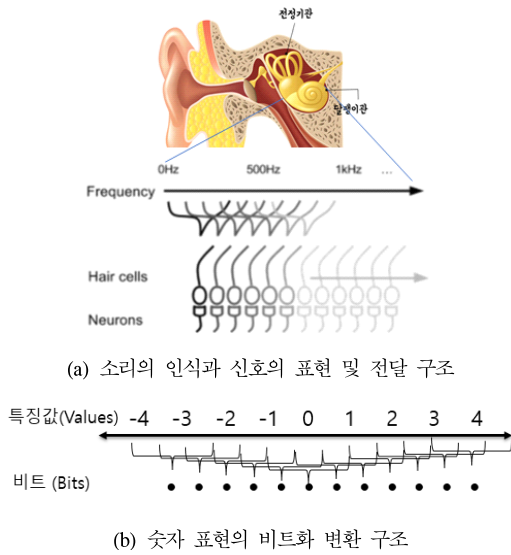


그림 5. 스칼라 인코딩 (scalar encoding) 개념도
Fig. 5. Concept diagram of scalar encoding

참고하면, 달팽이관의 유모세포(hair cell)는 소리의 주파수를 기반으로 한 일련의 뉴런을 자극한다. 이러한 자극은 중첩된 분포로 나타나며 특정 범위의 주파수 응답을 보이는 영역은 뉴런에 연결되어 음파 신호가 입력될 때 발화되는 구조로 그림 5의 (a)와 같이 되어 있다.

시각 인식으로 확장한다면, 이미지가 매니폴드 변환으로 부동소수점의 저차원 특징값으로 나타날 때 이를 표현하고 전달하기 위해서는 특정 뉴런을 자극하는 구조 및 최소 분포하는 뉴런의 발화와 유사하게 인코딩할 필요가 있다. 이를 위한 방법론으로 그림 5의 (b)와 같이 숫자를 중첩적으로 인코딩하는 방식을 적용하여 구현할 수 있다. 이 구조에서는 유사한 특징값을 가지는 영역은 비트스트림상에서 유사한 패턴으로 나타나고 이 패턴이 비트로 표현되기 때문에 AND 나 OR과 같은 논리 연산이 가능한 구조를 제공할 수 있다.

III. 시스템 구현 및 결과

3.1 개요

본 논문의 기억의 순환 메커니즘을 모델링 하기 위한 차원축소 및 특징값 인코딩 기반 가변형 이미지 표현시스템은 비교사 학습 기반으로 특징값을 생성하는 변분 오토인코더 기반 매니폴드 변환 모듈, 매니폴드 변환 모듈의 학습에서 필요한 하이퍼 파라미터를 튜

닝하기 위한 특징값 기반 분류기와 특징값 분석기/가시화 모듈, 저차원 변환된 부동소수점 기반 특징값을 비트스트림으로 인코딩하는 스칼라 인코더 모듈, 마지막으로 인코딩된 특징값을 표시하는 출력부로 그림 6과 같이 구성된다. 그림 6의 구성에서 매니폴드 변환 모듈의 디코더와 특징값 기반 분류기와 특징값 분석기/가시화 모듈은 신경망의 학습과정에서만 사용되고 매니폴드 변환 모듈의 인코더와 스칼라 인코더는 학습 및 실행과정 모두에서 사용된다.

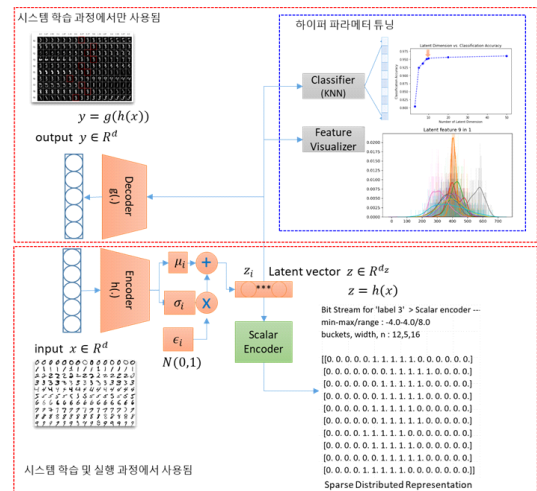


그림 6. 구현시스템 구성도
Fig. 6. Diagram of system implementation

3.2 변분 오토인코더 기반 매니폴드 변환 모듈

고차원 데이터를 저차원으로 변환하고 추상화하는 방법론으로, 신경망 학습 때 입력과 출력을 동일하게 적용하여 시스템 구조적으로는 유사하지만 사용되는 수학적 모델에 따라 분별(discriminative) 모델 기반의 오토인코더^[11]와 확률 모델을 근간으로 하는 생성(generative) 모델 기반의 변분 오토인코더^[12]를 사용할 수 있다. 분별 모델 기반 오토인코더는 학습 데이터가 충분히 많을 때 좋은 성능을 보이고 데이터를 분류/구분하는 목표를 두고 있어서 데이터가 실제 어떤 모습인지 본질을 이해하기는 어렵다. 반면에 생성 모델 기반 변분 오토인코더는 분별 모델 방식과는 다른 방식으로 분류 카테고리에 맞는 데이터를 확률 기반으로 추론하는 방식이다. 생성 모델에서는 생성되는 입력과 결과 데이터는 분류 클래스별로 특정한 통계적 분포를 따른다고 가정한다. 생성 모델의 장점은 학습 데이터셋의 규모가 적어도 적용할 수 있고 데이터 자체의 특성을 파악하기에 유리하다. 그리고 데이터를 생성할

수 있어 새로운 결과물을 얻어낼 수 있다. 그러나 데이터가 많을 때는 분류 성능 차원에서는 분별 모델과 비교하면 다소 성능이 떨어질 수 있다.

본 논문의 개념 검증용 구현에서는 구현 목적이 분류기가 아니라 학습에 사용하는 이미지 데이터의 속성을 모델링 하여 특징값으로 추출하는 것이므로 변분 오토인코더 모델이 적합하므로 기반 모델로 변분 오토인코더 모델을 활용하였다.

변분 오토인코더는 오토인코더의 특성을 물려받지만 다른 점이 있다. 오토인코더에서는 잠재변수(z)가 학습 데이터와 특별한 관련 없이 단순히 계산되는 중간값으로 결정론적(deterministic) 값일 뿐이지만 변분 오토인코더에서는 잠재변수(z)가 연속 분포를 가지는 랜덤 변수라는 점이 중요한 차이점이다. 즉 잠재 변수(z)가 평균(μ_i)과 표준편차(σ_i)로 결정되는 확률 분포를 가진다는 뜻이다.

변분 오토인코더의 인코더 신경망은 주어진 입력 데이터(x)로부터 잠재 변수(z)를 얻을 확률 p(z|x)로, 변분 오토인코더의 디코더 신경망은 잠재 변수(z)로부터 결과물(x, 출력을 입력과 동일하게 설정하여 학습)을 얻을 확률 p(x|z)로 나타낸다. 그림 7에서는 본 연구에서 구현한 매니폴드 변환 모듈의 내부 구조를 보여주고 있다.

그림 7에서 인코더 신경망은 주어진 입력(x)으로부터 평균(μ_i)과 표준편차(σ_i)를 가지는 가우시안 확률 분포를 학습하는 신경망이다. 본 실험 모델에서는 인코더 신경망으로 500개의 노드를 가지는 신경망 두 층을 완전연결(FCL, fully connected layer)하여 구현하였다.

최종적으로 잠재공간의 개별 잠재변수(z_i)는 평균(μ_i)과 표준편차(σ_i)의 가우시안 확률 분포에서 샘플링 된다. 샘플링 과정에서 도입된 변수 ϵ_i 는 학습에 의해 결정되는 것이 아니라 정규분포에서 샘플링되어 표준편차에 곱해져서 잠재변수를 만드는데 사용된다. 일반적으로 샘플링하는 분포에 대해 역전과 하는 것은 쉽지 않기 때문에 확률적 분포(stochastic) 노드를 확률적 분포(stochastic) 부분과 결정론적(deterministic) 부분으로 분해하여 결정론적 부분으로 역전과가 되도록 하는 재매개변수화 트릭(reparameterization trick)^[17]이 사용된다. 왜냐하면 재매개변수화 과정을 거쳐도 확률분포는 변경 전과 동일하기 때문이다.

변분 오토인코더의 디코더 신경망은 잠재변수(z)로부터 입력 데이터(x)의 확률 분포를 만들어내는 생성 모델 기반 신경망 네트워크이다. 본 실험 모델에서는 베르누이 확률 분포를 따르는 결과를 만들기 위한 디코더 신경망으로 500개의 노드를 가지는 신경망 두 층을 완전연결(FCL, fully connected layer)하여 구현하였다.

결과물 x(본 구현에서는 이미지임)의 확률 분포 p(x)를 알기 위해, 디코더 신경망은 p(x|z)를 학습한다. 생성 모델이므로 디코더 신경망은 잠재 변수(z)의 샘플 값이 달라짐에 따라 연속적으로 변하는 결과물(이미지)을 만들어낼 수 있다.

본 실험 모델 구현에서 그림 7의 디코더 블록은 잠재 공간(z)을 학습시키는 동안에만 사용된다. 그리고 신경망의 학습과정에서 사용된 내부 파라미터로 학습률(learning rate)은 0.001, 배치크기(batch size)는 100을 사용하였다.

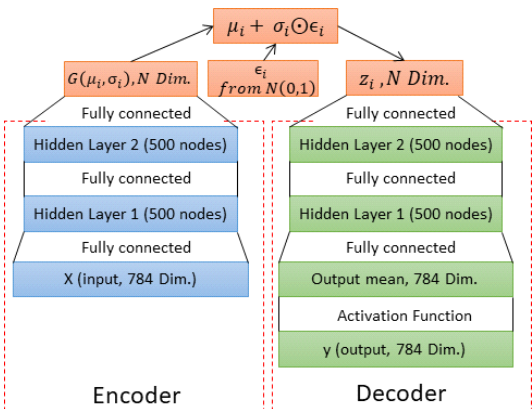


그림 7. 변분 오토인코더 기반 매니폴드 변환 모듈
Fig. 7. VAE based Manifold transform module

3.3 특징값 분석 기반 하이퍼 파라미터 튜닝 모듈
변분 오토인코더 기반 매니폴드 변환에서 중요한 중 하나는 비교사 학습을 위한 하이퍼 파라미터다. 내부적으로 변분 오토인코더는 인코더/디코더 신경망으로 구성되어 있으므로 신경망 내부 파라미터는 인코더/디코더를 이용한 학습으로 결정되지만 변분 오토인코더를 포함한 시스템 전반의 하이퍼 파라미터는 튜닝의 관점에서 결정해야 한다. 본 실험 구현에서 시스템 전반에 영향을 미치는 하이퍼 파라미터는 별도로 통계를 통해 결정하였으나 후속 연구를 통해 시스템 전반에 사용되는 파라미터를 자동화하는 방법을 같이 고려할 계획이다.

분류 성능 기준으로 결정한 잠재공간의 차원으로

입력 이미지가 차원축소되어 특징값들이 추출되었을 때, 추출된 이 특징값들의 분포를 확인해 볼 필요가 있어 본 개념 검증 구현에서는 직관적으로 특징값의 분포를 확인할 수 있도록 그래프 위주로 도식화하여 살펴보았다.

먼저 고차원의 이미지 데이터를 몇 차원의 데이터로 차원축소할 것인지 결정하기 위해 파이선 사이킷런(scikit-learn) 패키지에서 제공하는 k-최근접 이웃 분류기(k-NN)를 이용하여 분류 성능에서 번복점에 해당하는 지점 (본 실험에서는 95% 달성 지점)으로 잠재 벡터의 차원을 결정하였다. 이와 관련한 선행 연구^{16,7)} 결과를 참고하면 그림 8과 같이 잠재공간을 10차원으로 축소할 경우 분류기 관점에서 입력의 속성을 거의 모두 포함하는 것으로 확인되었다.

그림 9는 라벨 3번과 라벨 6번으로 분류되는 이미지에 대한 10개의 특징값(feature 0 ~ feature 9)을 하나의 그림에 누적하여 가시화한 것이다. 그림 9에서 그래프 각 라인의 색깔은 같은 특징값 번호를 의미하는데 예에서 보이는 바와 같이 입력데이터의 라벨에 따라 서로 다른 특징값 분포를 보이므로 매니폴드 변환이 수행되었음을 확인할 수 있다. 따라서 라벨에 따라 특징값을 누적하여 그린 이 그래프는 라벨에 따른 이미지를 대변하는 것으로 간주할 수 있다.

그림 10은 모든 라벨에 대한 직관적인 비교를 위해 라벨 0 ~ 라벨 9까지의 이미지에 대한 10개의 특징값을 도식하였고, 마지막 칼럼에는 10개의 특징값을 누적하여 하나의 그래프로 표시하였다. 다른 관점에서 살펴보면 마지막 칼럼의 그래프는 라벨에 따른 이미지를 대변하는 표현으로 간주할 수 있는데, 이 그래프의 실제 값은 부동소수점 숫자로 되어있다. 그러나 부동소수점으로 표시되는 특징값을 비트로 변환하면 추후 이미지의 검색 및 비교시 비트기반 논리 연산 등이

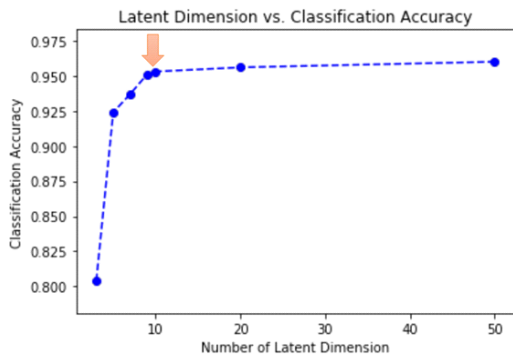


그림 8. 잠재공간 차원의 결정
Fig. 8. Decision of latent space dimension

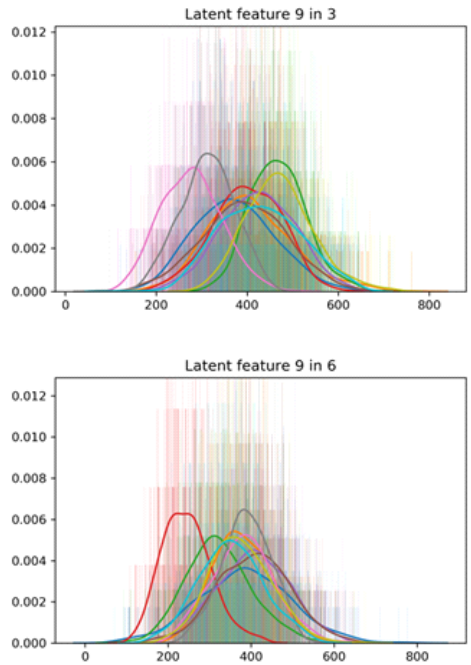


그림 9. 라벨3과 라벨6에 대한 특징값 누적의 예
Fig. 9. Example of Accumulation of Feature Values for Labels 3 and 6

가능한 구조를 제공할 수 있으므로 이를 비트스트림으로 변환할 필요가 있다.

3.4 스칼라 인코딩 모듈

부동소수점으로 표시된 이미지에 대한 특징값을 비트로 변환하면 비트의 논리연산을 통해 이미지를 빠르게 검색하거나 논리적 비교가 가능하므로 이미지에 대한 특징값을 비트스트림으로 표시할 필요가 있고 이를 위한 인코더가 필요하다.

통상적으로 데이터 인코더의 설계는 데이터 유형 및 애플리케이션에서 중요하게 다루어지는 데이터의 의미적 특성을 캡처해야 한다. 데이터를 비트스트림으로 인코딩할 때 고려해야 할 몇 가지 중요한 측면으로는 먼저 의미론적으로 유사한 데이터는 서로 중복되는 비트를 가져야 하고, 동일한 입력에 대해서는 동일한 출력 비트스트림을 생성해야 한다. 그리고 출력은 모든 입력에 대해 동일한 차원 (전체 비트 수)를 가져야 한다.

본 실험 구현에서는 부동소수점 숫자로 표현되는 특징값을 상술한 데이터 인코더 특징을 만족하는 인코딩을 위해 스칼라 인코더를 활용하여 특징값을 비트스트림으로 변환하였고 그 절차는 아래와 같다.

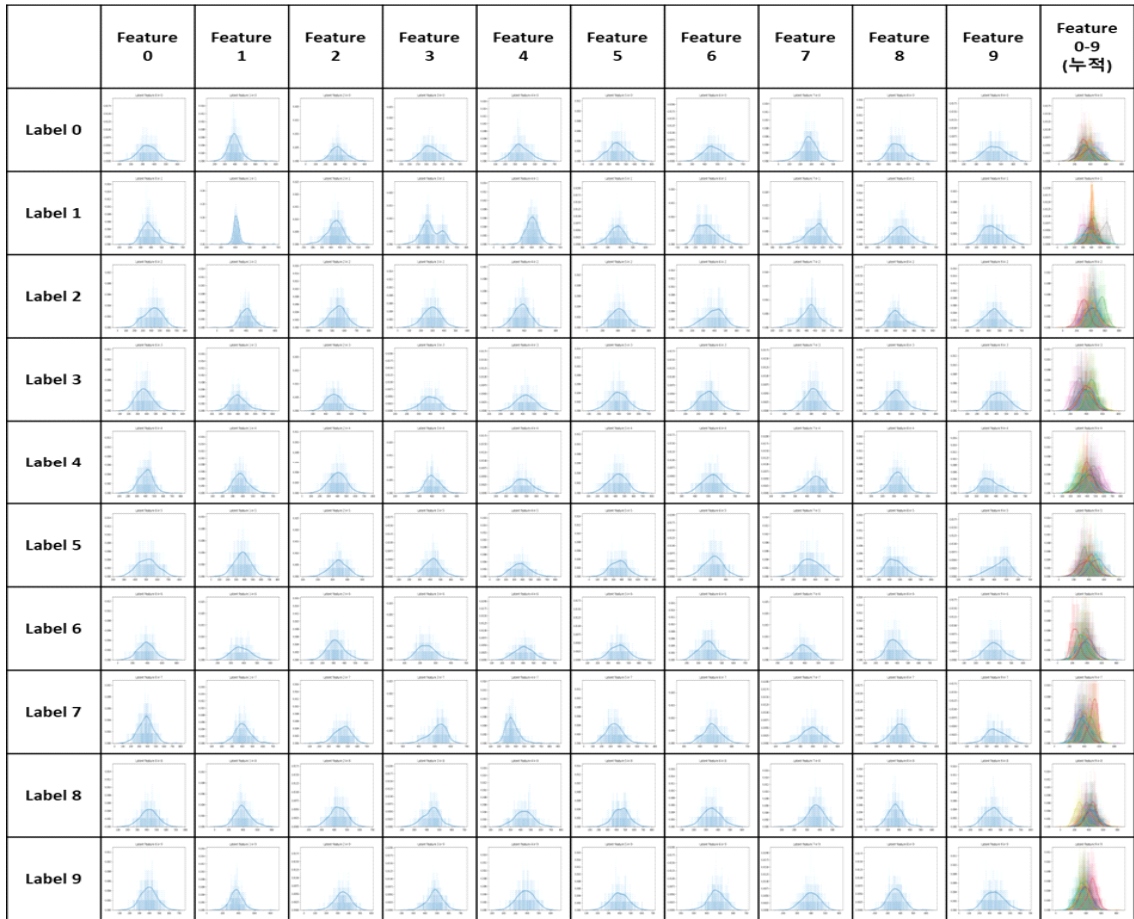


그림. 10. 이미지 라벨의 저차원 변환된 특징값의 분포 및 시각화 예시
 Fig. 10. Example of distribution and visualization of low-dimensional transformed feature values of an image label

- 특징값이 나타내는 값의 범위를 minVal, maxVal 값으로 선택 (minVal, maxVal): 변분 오토인코더 모듈의 출력값 범위는 최대 4.0 최소 -4.0의 범위를 가지므로 maxVal = 4.0, minVal = -4.0로 설정됨
- 표현하고자 하는 값의 범위를 계산 (range = maxVal - minVal): Range = 8.0으로 설정됨
- 특징값을 몇 개로 분할할 것인지 선택 (number of buckets, buckets): 출력 스케일에 따라 가변적으로 계산됨
- 인코딩 완료된 비트스트림 표현에서 사용할 활성 비트 수를 결정 (width, w): 출력 스케일에 따라 가변적으로 계산됨
- 인코딩에 사용할 총 비트수 계산 (number of bits, n = buckets + w - 1): 이미지의 특징값을 몇 비트로 표현할지 사용자가 선택

- 특징값으로 주어진 부동소수점 값(v, value)은 i번째 버킷(i, buckets)으로 결정됨 ($i = \text{floor}[\text{buckets} * (v - \text{minVal}) / \text{range}]$)

본 논문의 개념 검증 구현에서는 이미지가 10차원의 특징값으로 변환되기 때문에, 각 이미지당 10개의 특징값을 가지므로 총 10개의 이미지 속성을 상호 중첩하여 하나의 비트스트림으로 이미지를 표현하는 구조로 생각할 수 있다.

그림 11은 이러한 변환 과정의 간단한 설명을 위해 단순화해서 표현한 것으로 10차원 특징값을 각 특징값별 16 비트로 변환한 예로 총 비트수는 160 비트(16 bits x 10 features)로 표현할 수 있다. 여기서 라벨 ‘3’의 특징값은 최소값(minVal)이 -4.0이고 최대값(maxVal)이 4.0의 범위(range)를 가지는 숫자이고 비트스트림에서 활성비트(width)가 5개가 되도록 하면,

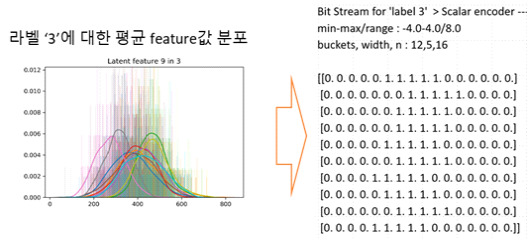


그림 11. 라벨 '3'에 대한 특징값 인코딩의 예
Fig. 11. Example of Feature Value Encoding for label '3'

각 특징값을 16 비트로 표현할 때 서로 다른 표현으로 구분되는 버킷(bucket)의 수는 12로 계산된다. 따라서 이 경우에는 각 특징값을 12 단계로 매핑하여 표현한 것이라 할 수 있다.

이미지 특징값을 그림 11과 같이 비트스트림으로 표현하게 되면 이미지에 대한 AND나 OR 등의 논리 연산을 통해 이미지 검색이 가능한 구조를 제공할 수 있다.

IV. 실험 결과 및 고찰

4.1 실험 환경 구성

본 논문의 개념 검증 구현의 실험 환경은 다음과 같이 구성하였다. 첫 번째 입력데이터는 다양한 화상 처리 시스템에서 많이 사용하는 MNIST 데이터셋을 이용하였다. MNIST 데이터셋은 손으로 쓴 '0'에서 '9'까지의 숫자들로 이루어진 28x28 크기의 이미지 데이터베이스이며, 다양한 화상 처리 시스템을 학습하기 위해 일반적으로 많이 사용된다. MNIST는 60,000 개의 훈련셋과 10,000개의 테스트셋으로 이루어져 있고 이중 훈련셋을 학습데이터로 사용하고 테스트셋은 신경망을 검증하는 데 사용한다. 두 번째 하이퍼 파라미터 학습을 위한 인공지능망은 변분 오토인코더 기반 매니폴드 변환을 사용하였다. 세 번째 컴퓨팅 환경으로는 GPU(TITAN RTX)가 내장된 윈도우 10 설치 데스크톱 컴퓨터에 아니콘다 실험환경을 설치하고 파이썬 3, 텐스플로를 설치하여 학습 및 실험을 시행하였다.

4.2 실험 결과 및 고찰

본 절에서는 실제 입력데이터인 MNIST 이미지를 입력으로 하여 변분 오토인코더 기반 매니폴드 변환을 거쳐 만들어진 저차원 특징값과 이를 비트스트림으로 인코딩한 결과를 예시하고 그 결과에 대해 살펴본다.

4.2.1 입력 이미지와 변환된 이미지의 비교

그림 12는 784 차원의 입력 이미지와 기 학습된 변분 오토인코더에 의해 10차원으로 차원 축소된 특징값으로부터 복원한 영상, 차원 축소된 특징값을 16/32/256비트 스칼라 인코딩을 거친 값을 역변환한 후 복원한 영상을 비교하였다.

변분 오토인코더에 의해 차원 축소된 특징값은 입력 이미지의 확률 분포와 가장 근접하는 평균과 표준편차를 학습하고 이를 기반으로 이미지를 복원하기 때문에 기본적으로 블러링이 발생한다. 따라서 그림 12의 (1)과 (2)의 비교에서 (2)는 논문에서 사용한 기본 변분 오토인코더의 최대 복원 성능이다. 본 논문의 목표는 변분 오토인코더의 신경망 구조 개선에 초점을 맞춘 것이 아니었기 때문에 이에 대한 연구는 후속 연구로 진행될 예정이다.

그림 12에서 (3), (4) 및 (5)는 서로 다른 비트로 스칼라 인코딩된 특징값을 역변환하여 이미지로 복원한 영상인데 (3), (4) 및 (5)의 최대 성능은 (2)가 된다. (3), (4) 및 (5)의 변환에서 특징값을 표현하는 총 비트는 서로 상이하고 활성비트('1')의 개수는 다섯개 (스칼라 인코딩의 width = 5)가 되도록 했다. 이렇게 설정할 경우 16비트로 변환 때 서로 다른 값을 표현할 수 있는 버킷수는 12개, 32비트로 변환 때 28개, 256비트로 변환 때 252개가 된다. 특징값의 범위가 8 (스칼라 인코딩의 range = 8)이므로 이산적으로 표현할 경우 특징값이 양자화 되는 간격은 16비트변환의 경우 8/12 (0.667), 32비트 변환의 경우 8/28 (0.29), 256비트 변환의 경우 8/252 (0.03)가 된다. 즉 표현되는 비트수가 작을수록 양자화 에러를 많이 포함하여 원

(1) 원본 입력 이미지 (784 dim.)	(2) VAE 변환 후 복원한 이미지 (10 dim.)	(3) 16 bit 변환 후 복원한 이미지 (12 buckets)	(4) 32 bit 변환 후 복원한 이미지 (28 buckets)	(5) 256 bit 변환 후 복원한 이미지 (252 buckets)

그림 12. 입력 이미지와 변환된 이미지 비교 예시
Fig. 12. Example of comparing input images with converted images

본과의 차이가 벌어진다. 그러나 이미지로 복원한 결과 그림 12의 (3), (4) 및 (5)는 (2)와 큰 차이를 보이지 않는다. 이는 변분 오토인코더가 입력 이미지의 특징값을 잘 클러스터링하고 있다는 것을 의미한다.

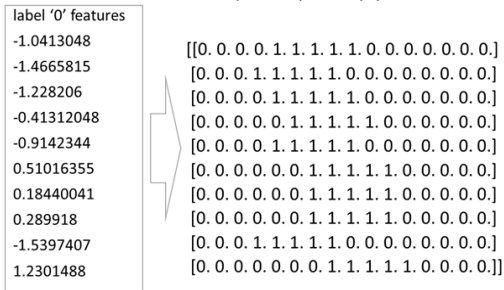
4.2.2 이미지를 비트스트림으로 변환한 결과

그림 13은 라벨 ‘0’ 및 라벨 ‘9’의 이미지를 저차원 변환하고 추출한 10개의 특징값을 각각 예시하고 각 특징값을 16비트로 인코딩 및 중첩한 비트스트림을 나타내었다.

그림 14는 라벨 ‘0’ 이미지를 저차원 변환하고 추출한 10개의 특징값을 32비트로 인코딩한 비트스트림의 예이다.

그림 13 및 그림 14에서와같이 동일한 라벨의 이미지를 서로 다른 스케일로 표현할 수 있는데 이는 서로

- Bit Stream for 'label 0' > Scalar encoder -----
- min-max/range : -4.0-4.0/8.0
- buckets, width, n : 12.5,16



(a) 라벨 ‘0’의 특징값 및 비트스트림

- Bit Stream for 'label 9' > Scalar encoder -----
- min-max/range : -4.0-4.0/8.0
- buckets, width, n : 12.5,16



(b) 라벨 ‘9’의 특징값 및 비트스트림

그림 13. 특징값과 변환된 비트스트림 예 (라벨 0과 9의 16비트 표현)
Fig. 13. Example of feature values and bitstreams (label 0 and 9 with 16 bits resolution)

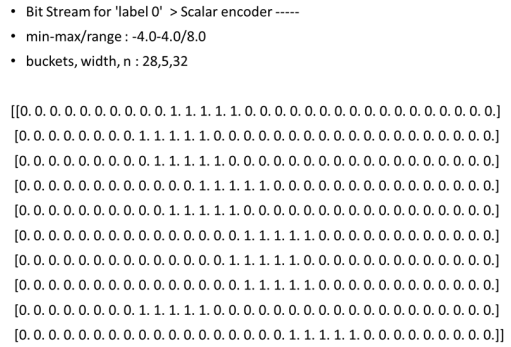


그림 14. 특징값과 비트스트림 예(라벨 0의 32비트 표현)
Fig. 14. Example of feature values and bitstreams (label 0 with 32 bits resolution)

다른 해상도로 이미지를 비트스트림으로 표현할 수 있다는 것을 의미한다.

4.2.3 특징값이 복원 이미지에 미치는 영향

각각의 특징값 들이 실제 복원 영상에서 어떤 영향을 미치는지 살펴보기 위해 라벨 ‘3’의 이미지에서 추출한 10개의 특징값(f0 ~ f9) 중에서 하나의 특징값만 -4.0에서 4.0까지 가변시키고 나머지 9개의 특징값은 그대로 사용하여 영상을 복원한 예를 그림 15에 나타내었다. 즉 그림 15의 가로 방향 행은 가변시키는 특징값 번호를 명기한 것이고 세로 방향 열은 가변되는 특징값을 표기한 것이다. 예를 들어 첫 번째 행은 첫 번째 특징값(f0)을 -4.0에서 4.0까지 가변시키며 복원한 라벨 ‘3’의 이미지로 여기서 빨간색 박스로 표시된 부분은 실제 계산된 특징값(f0)이 적용된 부분이다.

그림 15에서 살펴보면, 라벨 ‘3’에 대해 만들어진 특징값 8번(f8)의 경우 특징값의 변화에 대해 복원 영상이 붉은 ‘8’에서 가늘어지는 ‘3’으로 변형을 유발한

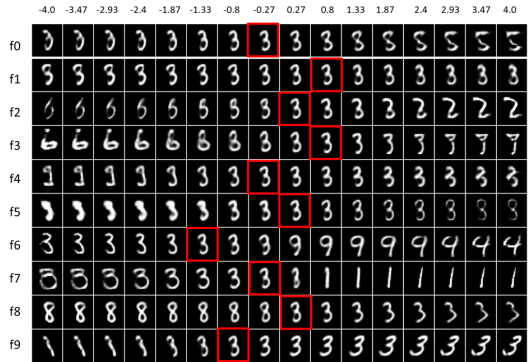


그림 15. 라벨 ‘3’의 특징값 변화에 따른 복원 영상의 비교
Fig. 15. Comparison of label ‘3’ reconstructed images according to changes in feature values

다는 것을 확인할 수 있다. 다른 관점에서 살펴보면, 그림 15에서 빨간색 박스로 표시된 ‘3’번은 실제 계산된 특징값을 모두 적용하여 복원한 영상으로, 차원 축소된 공간에서 원본데이터 라벨 ‘3’의 이미지에 대한 대표 표현으로 간주할 수 있다. 달리 말하면 실세계 원본데이터 라벨 ‘3’을 만들려면 $f_0 \sim f_9$ 의 특징값들이 상호 얼마씩 섞여야 하는지를 직관적으로 보여준다.

다른 관점에서 몇 가지 특징값을 살펴보면, 특징값 5번(f_5)은 이미지에서 라인의 굵기를 결정하는 속성으로 간주할 수 있고 특징값 9번(f_9)은 이미지에서 라인의 기울기를 나타내는 속성이라고 간주할 수 있다. 그러나 어떤 특징값이 이미지에 대한 어떤 속성을 나타낸다고 명명하기는 아직 부적절하며 이에 관한 추가적인 연구 또한 필요하다.

본 논문에서 제안한 시스템은 개념 검증 연구의 실험결과로 몇 가지 한계점이 있다. 따라서 향후 보완 연구에서는 몇 가지 사항에 관한 연구가 추가되어야 하는데 요약하면 아래와 같다.

- 개념 검증 관점의 구현이라 변분 오토인코더의 인코더/디코더 부분을 복잡하게 다루지 않고 간단한 구조로 하였으므로 입력데이터에 따라 하이퍼 파라미터를 가변적으로 변형 적용할 필요가 있음
- 개념 검증 구현에서는 세부적 특징값에 대한 상관도 분석 대신 직관적으로 차이를 눈으로 확인할 수 있도록 그래프로 위주로 살펴보았으나 향후 정량적 유사도 분석을 병행할 필요가 있음 - 평균제곱차이 유사도 (Mean Squared Difference Similarity), 코사인 유사도 (Cosine Similarity), 피어슨 유사도 (Pearson Similarity) 등을 고려
- 고차원의 이미지 데이터를 몇 차원의 데이터로 차원축소할 것인지 결정하기 위해 기본적인 k-NN을 이용하여 분류 성능이 95%를 달성할 수 있는 수준으로 잠재벡터의 차원을 결정하였으나 상황에 따라 하이퍼 파라미터를 동적으로 결정할 방법론을 결정하기 위한 연구가 필요함
- 특징값으로 나타나는 속성을 네이밍하고 정량화 할 방안에 대한 후속 연구가 필요함

V. 결 론

본 논문은 기억의 순환 메커니즘을 모델링하기 위한 차원축소 및 특징값 인코딩을 통한 가변형 이미지 표현방법에 관한 연구로 매니폴드 변환 기반으로 입력데이터의 차원을 축소하고 추출한 특징값을 활용하

여 비트스트림으로 인코딩하여 이미지를 가변적으로 표현하는 개념 검증 연구이다. 본 연구에서 차원축소 및 특징값 추출을 위한 매니폴드 변환 기법으로는 변분 오토인코더를 활용하였고, 특징값을 인코딩하기 위해서는 스칼라 인코딩을 응용하였다.

본 연구의 가용성을 확인하기 위해 MNIST 데이터셋을 활용하여 제안한 방법론의 가용성을 확인하는 개념 검증 실험을 수행하였다. 개념 검증 실험에서는 고차원의 MNIST 이미지가 매니폴드 변환에 의해 10차원으로 축소되고, 축소된 10차원 특징값을 스칼라 인코딩을 통해 가변 비트의 비트스트림으로 변환이 가능함을 확인하였다.

본 연구는 개념 검증 연구 단계라 다양한 입력에 대한 테스트를 시행하지 못했지만, 제한적인 범위에서 이미지를 논리 연산이 가능한 비트스트림으로 표현 가능하다는 것을 확인하였다.

향후 제안한 시스템의 성능 향상을 위해 추가로 개발이 진행되어야 할 부분은 첫 번째로 다양한 데이터셋에 대한 실험을 실시하여 원본 데이터의 정보량 손실 제어가 가능한 구조의 비트스트림으로 변환할 수 있도록 하는 추가 연구가 필요하고, 두 번째는 잠재공간을 비트스트림으로 변환하는 과정에서 비트의 공간 분포를 확률적으로 변경할 수 있는 해시함수 같은 변환 함수를 도입할 필요가 있다.

References

- [1] Defense Advanced Research Projects Agency (DARPA), *Explainable Artificial Intelligence (XAI)*, DARPA presentation, Retrieved 24, Apr. 2020, from <https://www.darpa.mil/>
- [2] H. Song, H. Kim, E. Jeong, S. Oh, J. Lee, D. Kang, J. Jeong, and Y. Lee, “Self-improving artificial intelligence technology,” *Electron. and Telecommun. Trends*, vol. 34 no. 4, pp. 43-54, 2019.
- [3] D. P. Kingma and M. Welling, *Auto-encoding variational bayes*, 2013, Retrieved Feb. 29, 2020, from <https://arxiv.org/abs/1312.6114>.
- [4] S. Purdy, *Encoding data for HTM systems*, pp. 1-11, 2016, Retrieved Feb. 29, 2020 from <https://arxiv.org/abs/1602.05925>.
- [5] L. Deng, “The MNIST database of handwritten digit images for machine learning research,” *IEEE Sign. Process. Mag.*, vol. 29,

no. 6, pp. 141-142, 2012.

[6] K. Kang and J. Cho, "Case study on category classification using manifold learning-based features," in *Proc. KICS Autumn Conf.*, pp. 602-603, Nov. 2019.

[7] K. Kang and J. Cho, "Feature analysis on manifold transformed space," in *Proc. KICS Winter Conf.*, Feb. 2020.

[8] Z. Deng, X. Zhu, D. Cheng, M. Zong, and S. Zhang, "Efficient kNN classification algorithm for big data," *Neurocomputing*, vol. 195, pp. 143-148, 2016.

[9] K. Yun, S. Kwon, K. Kim, S. Kim, J. Nam, S. Seo, H. Song, M. Sin, S. You, Y. Lee, H. Lee, W. Jeon, and S. Chun, "Understanding psychology," *Hajisa*, Jun. 2019.

[10] Y. Bengio, A. Courville, and P. Vincent, "Representation learning: a review and new perspectives," *IEEE Trans. Pattern Anal. and Machine Intell.*, vol. 35, no. 8, pp. 1798-1828, 2013.

[11] A. Ng, *Sparse autoencoder*, CS294A Lecture Notes, p. 72, 2011, Retrieved Feb., 2020, from https://nlp.stanford.edu/~socherr/sparse-Autoencoder_2011new.pdf.

[12] J. B. Tenenbaum, V. Silva, and J. C. Langford, "A global geometric framework for nonlinear dimensionality reduction," *Science*, vol. 290, no. 5500, pp. 2319-2323, 2000.

[13] S. T. Roweis and L. K. Saul, "Nonlinear dimensionality reduction by locally linear embedding," *Science*, vol. 290, no. 5500, pp. 2323-2326, 2000.

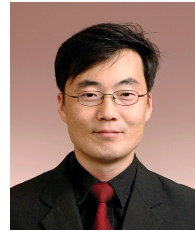
[14] L. Maaten and G. Hinton, "Visualizing data using t-SNE," *J. Machine Learning Res.*, pp. 2579-2605, 2008.

[15] Wikipedia, *Two-streams hypothesis*, Retrieved Feb. 29, 2020, from https://en.wikipedia.org/wiki/Two-streams_hypothesis.

[16] J. Hawkins and S. Blakeslee, "On intelligence: How a new understanding of the brain will lead to the creation of truly intelligent machines," Macmillan, 2007.

[17] D. P. Kingma, T. Salimans, and M. Welling, "Variational dropout and the local reparameterization trick," in *Advances in NIPS*, pp. 2575-2583, 2015.

강 규 창 (Kyuchang Kang)



2009년 8월: 충남대학교 컴퓨터공학과 공학박사
 1997년~2001년: 국방과학연구소 종합시험단 연구원
 2001년~2017년: 한국전자통신연구원 SW콘텐츠연구소 책임연구원

2017년~현재: 군산대학교 IT정보제어공학부 (IT융합통신공학전공) 조교수
 <관심분야> 인공지능, 시각지능, 개방형 소프트웨어 플랫폼, 사물인터넷
 [ORCID:0000-0003-0833-8906]

조 주 필 (Juphil Cho)



2001년 2월: 전북대학교 전자공학과 공학박사
 2000년~2005년: 한국전자통신연구원 이동통신연구단 선임연구원
 2006년~2007년: 한국전자통신연구원 이동통신연구단 초빙연구원

2011년~2012년: 미국 NSF, 교환교수
 2005년~현재: 군산대학교 IT정보제어공학부 (IT융합통신공학전공) 교수
 <관심분야> LTE-A, 5G 이동통신, 차세대 WLAN, Cognitive Radio, LED-ID, Smart Farm, Smart IoT, 방송통신융합기술
 [ORCID:0000-0003-1041-2538]