

다중 생체신호를 이용한 신경망 기반 전산화 감정해석

Neural-network based Computerized Emotion Analysis using Multiple Biological Signals

이지은* · 김병남* · 유선국*†

Jee-Eun Lee* · Byeong-Nam Kim* · Sun-Kook Yoo*†

*연세대학교 의과대학 의학공학교실

*Department of Medical Engineering, Yonsei University College of Medicine

Abstract

Emotion affects many parts of human life such as learning ability, behavior and judgment. It is important to understand human nature. Emotion can only be inferred from facial expressions or gestures, what it actually is. In particular, emotion is difficult to classify not only because individuals feel differently about emotion but also because visually induced emotion does not sustain during whole testing period. To solve the problem, we acquired bio-signals and extracted features from those signals, which offer objective information about emotion stimulus. The emotion pattern classifier was composed of unsupervised learning algorithm with hidden nodes and feature vectors. Restricted Boltzmann machine (RBM) based on probability estimation was used in the unsupervised learning and maps emotion features to transformed dimensions. The emotion was characterized by non-linear classifiers with hidden nodes of a multi layer neural network, named deep belief network (DBN). The accuracy of DBN (about 94 %) was better than that of back-propagation neural network (about 40 %). The DBN showed good performance as the emotion pattern classifier.

Key words: Emotion, Biological Signal, Restricted Boltzmann Machine (RBM), Multilayer Neural Network (MNN), Deep Belief Network (DBN)

요약

감정은 학습능력, 행동, 판단력 등 삶의 많은 부분에 영향을 끼치므로 인간의 본질을 이해하는 데 중요한 역할을 한다. 그러나 감정은 개인이 느끼는 강도가 다르며, 시각 영상 자극을 통해 감정을 유도하는 경우 감정이 지속적으로 유지되지 않는다. 이러한 문제점을 극복하기 위하여 총 4가지 감정자극(행복, 슬픔, 공포, 보통) 시 생체신호(뇌전도, 맥파, 피부전도도, 피부 온도)를 획득하고, 이로부터 특징을 추출하여 분류기의 입력으로 사용하였다. 감정 패턴을 확률적으로 해석하여 다른 공간으로 매핑시켜주는 역할을 하는 Restricted Boltzmann Machine (RBM)과 Multilayer Neural Network (MNN)의 은닉층 노드를 이용하여 비선형적인 성질의 감정을 구별하는 Deep Belief Network (DBN) 감정 패턴 분류기를 설계하였다. 그 결과, DBN의 정확도(약 94%)는 오류 역전과 알고리즘의 정확도(약 40%)보다 높은 정확도를 가지며 감정 패턴 분류기로서 우수성을 가짐을 확인하였다. 이는 향후 인지과학 및 HCI 분야 등에서 활용 가능할 것으로 사료된다.

주제어: 감정, 생체신호, Restricted Boltzmann Machine (RBM), Multilayer Neural Network (MNN), Deep Belief Network (DBN)

※ 본 연구는 연세대학교 의과대학 교내연구비(6-2014-0010)의 지원에 의하여 이루어진 연구로서, 관계부처에 감사드립니다.

† 교신저자 : 유선국 (연세대학교 의과대학 의학공학교실)

E-mail : SUNKYOO@yuhs.ac

TEL : 02-2228-1919

FAX : 02-363-9923

1. 서론

감정의 분류를 위해서는 감정의 정의를 아는 것이 중요하다. Kleinginna(1985)는 신경과 호르몬 시스템으로부터 제어되는 주관적이고 객관적인 요소들의 복합적 상호작용을 감정이라고 정의했다. 감정은 인간의 삶에서 많은 역할을 하고 있으며, 어떠한 감정을 느끼는지에 따라 인간의 행동, 판단력, 학습능력 등이 결정된다. 따라서 감정을 정확히 판별하는 것은 상호간 의사소통 능력을 향상시키고 스트레스 유발을 개선하며, 경우에 따라 건강에 영향을 미치는 감정인자를 찾고 예방할 수 있다.

감정은 대표적으로 2가지 방법으로 분류하고 있다. 하나는 6가지 감정인 행복, 슬픔, 공포, 놀람, 역겨움, 분노를 기본으로 다른 감정들을 이러한 기본 감정의 한 부분으로 분류하는 것이고, 다른 하나는 Valence-Arousal 이차원 평면에 매칭하는 것이다(Krause, 1987; Lang, 1995).

감정을 유발 시킨 후에 리커트 척도(Likert Scale)와 자기설문평가(Self Assessment)를 통하여 감정을 유추할 수 있으나 설문조사는 피험자의 주관적 생각을 반영하는 한계점을 가진다(Niu et al., 2011). 인간이 인지하는 모든 것들은 생체 내 반응으로 나타나게 되며 생체신호를 통하여 확인할 수 있다. 외부에서 생체에 자극을 주게 되면 중추신경계와 자율신경계를 통하여 반응이 나타나게 된다(Guang-yuan & Min, 2009). 자극에 따른 신경계의 반응을 통하여 측정할 수 있는 생체신호로는 뇌파, 심전도, 맥파, 피부전도도, 피부 온도, 호흡, 근전도 등이 있다(Guang-yuan & Min, 2009). 이러한 생체신호들은 감정 판별에 객관적 정보로 이용할 수 있어 주관적인 감정 판별의 한계를 극복할 수 있다.

감정을 분류하기 위해 기존에는 서포트 벡터 머신(Support Vector Machine, SVM), 신경회로망(Neural Network, NN) 등의 교사학습 및 K-means를 이용한 비교사학습 등 다양한 기법이 사용되었다(Jerritta et al., 2011; Khashman, 2008; Murugappan et al., 2010; Peng et al., 2014; Schaaff & Schultz, 2009; Wang & Shang, 2013). 그러나 감정은 개인에 따라 느끼는 강도가 다르며, 시각 영상 자극을 통해 감정을 유도하는 경우

같은 감정이 지속적으로 유지되지 않으므로, 영상에 따른 라벨링을 통한 분류기 모델링은 감정 판별 분류 정확도가 떨어진다.

이에 따라 본 논문에서는 시각 영상 자극을 활용하는 경우, 감정 유발 및 판별의 정확도가 떨어지는 한계가 있기 때문에 이를 해결하기 위하여 생체신호를 이용하여 교사학습 및 확률적 해석과 비교사학습이 이루어지는 Deep Belief Network (DBN) 모델링을 통해 보다 높은 정확도를 가지는 향상된 감정 상태 분류기를 설계하였다.

2. 본론

2.1. 데이터 획득

본 연구에서는 시각자극인 비디오 영상을 통하여 행복, 슬픔, 공포, 보통 시의 4가지 감정상태에 대한 생체신호를 취득하였다. 사용된 감정유발 영상은 각각 ‘맘마미아(mamma mia)’, ‘폴빵엄마’, ‘쏘우3 (saw3)’, ‘역사스페셜’이다. 피험자는 아무런 자극이 없는 상태로 30분간 휴식을 취한 후 진행되었다. 실험은 총 75분으로 구성되었으며, 처음 15분 동안은 다큐멘터리 영상을 시청하도록 하여 참조 값을 측정하였고 나머지 60분 동안은 감정유발 영상을 시청하도록 진행되었다. 총 4개의 감정유발 영상을 하루에 한 개씩 시청하여 총 4일 간 실험을 진행하였다(Choi, 2011).

피험자는 건강에 이상이 없는 평균 60세의 노년층 여성 10명을 대상으로 진행되었다. 피험자의 생체신호는 BIOPAC MP 150TM을 이용하여 측정하였으며 1 kHz로 샘플링 한 뇌파, 맥파, 피부전도도 및 피부 온도 데이터를 획득하였다. 뇌파는 10-20 시스템을 기준으로 하여 전두엽(F3, F4)과 두정엽(C3, C4)에서 측정되었으며, 기준전극으로 피험자의 왼쪽 귀를 사용하였다. 맥파는 왼손 검지, 피부전도도는 왼손 중지와 약지, 피부온도는 왼손 소지에서 측정되었다(Choi, 2011).

2.2. 특징 추출

본 논문에서는 뇌파, 맥파, 피부전도도, 피부 온도

를 기반으로 하는 특징을 입력으로 사용하기 위하여 측정된 생체신호로부터 특징벡터를 추출했다.

자극유발 영상 시작 후 초기 5분 데이터는 감정의 유발이 일어나지 않는다는 가정 하에 이를 제외한 55분 데이터에서 특징을 추출했다. 특징은 획득 데이터에 5분 기반 창(Window)을 사용하여 계산하였으며 30초씩 창 이동(Sliding Window)을 하였다. 이에 따라 한 개의 감정자극 실험에서 100개의 데이터를 얻을 수 있었으며, 피험자 1명당 400개의 데이터 세트를 추출하였다.

뇌파에서의 특징추출을 위해 데이터는 Equation 1을 통하여 주파수 영역으로 표현한다. 여기서 N 은 샘플 개수, x 는 시간 영역의 신호, X 는 주파수 영역에서의 각 주파수에 해당하는 푸리에 변환(Fourier Transform) 결과를 나타낸다. Equation 2는 푸리에 변환으로부터 구한 X_k 의 계수를 이용하여 뇌파의 구성 주파수 밴드별 파워인 P_k 를 계산하여 뇌파의 특징을 추출하는 방법이다(Moretti et al., 2004). 추출된 δ , θ , α , β , γ 파의 주파수 파워 중 δ 파의 주파수 파워는 수면상태 혹은 잠음을 나타내고, γ 파의 주파수 파워 역시 고주파 성분이 많이 들어있으므로 특징에서 제외하고 θ , α , β 파의 주파수 파워를 특징으로 선택하여 주었다(Murugappan et al., 2010). 따라서 뇌파에서는 전두엽(F3, F4)과 두정엽(C3, C4) 총 네 군데 영역에서 각 3개의 주파수 파워 특징을 추출하여 총 12개의 특징벡터를 사용하였다.

$$X_k = \sum_{n=0}^{N-1} x_n e^{-\frac{j2\pi}{N}kn} \quad k=0,1,\dots,N-1 \quad (1)$$

$$P_k = \sum_{i=1}^N X_k X_k^* \quad (2)$$

맥파에서 추출하고자 하는 특징은 심전도의 특징, 맥파 전달시간(Pulse Transit Time) 및 호흡에 관한 특징이다. 먼저 심전도의 특징은 심박변이도로부터 추출할 수 있다. 심박변이도는 심박동 주기 사이의 변화를 나타내며 R 피크 간 간격을 통하여 구할 수 있다. 심박변이도를 통하여 구할 수 있는 시간 영역 특징은 심박변이도의 평균, SDNN, RMSSD, NN50이며 Equation 3, Equation 4, Equation 5 그리고 Equation 6을 통하여 구할 수 있다. $R \sim R$ 은 R 피크 사이의 간

격을 나타내며 $the\#of\ R \sim R$ 은 R 피크 간격의 개수를 나타낸다. 심박변이도의 주파수 영역 특징은 뇌파 특징 추출 시와 같은 방법으로 주파수 파워를 구할 수 있으며 고주파파워와 저주파파워의 비(LF/HF)를 특징으로 사용하였다(Malik & Camm, 1990). 심박변이도의 특징은 감정 관련 연구에서 일반적으로 사용되고 있는 분석방법이며, 본 연구에서 사용하는 심박변이도 특징은 총 5개이다.

$$MEAN_{HRV} = \frac{\sum(R \sim R)}{the\#of\ R \sim R} \quad (3)$$

$$SDNN = \sqrt{\frac{\sum\{(R \sim R) - MEAN_{HRV}\}^2}{the\#of\ R \sim R}} \quad (4)$$

$$RMSSD = \sqrt{\frac{\sum_{i=1}^{the\#of\ R \sim R} (R \sim R_{i+1} - R \sim R_i)^2}{the\#of\ R \sim R}} \quad (5)$$

$$NN50 = the\#of\ (\sum_{i=1}^{the\#of\ R \sim R} (R \sim R_{i+1} - R \sim R_i) > 50ms) \quad (6)$$

$$If, R \sim R_{i+1} - R \sim R_i > 50ms$$

또한, 맥파의 간접 유도를 통하여 맥파 전달시간 특징을 추출했다. 맥파 신호를 통한 R 피크 및 맥파 주기의 최하점을 이용하여 혈류량의 시차를 맥파 전달시간으로 정의하며 맥파 전달시간의 평균과 표준편차를 특징으로 선택하였다. 맥파 전달시간의 평균은 Equation 7을 이용하여 계산하며 $R \sim P$ 는 R 피크와 맥파의 최하점 사이의 시간 간격을 나타내며 $the\#of\ R \sim P$ 는 간격들의 개수를 나타낸다. 표준편차는 SD로 나타내었으며 Equation 8을 통하여 구하였다(Malik & Camm, 1990). 맥파로부터 추출 가능한 호흡신호는 심박변이도 주파수 특징과 유사한 범위를 가지기 때문에 사용하지 않았다. 따라서 맥파에서 추출한 특징은 총 7개로 심박변이도와 맥파 전달시간과 관련된 특징 값이다.

$$MEAN_{PTT} = \frac{1}{the\#of\ R \sim P} \sum(R \sim P) \quad (7)$$

$$SD_{PTT} = \sqrt{\frac{1}{the\#of\ R \sim P - 1} \sum\{(R \sim P) - MEAN_{PTT}\}^2} \quad (8)$$

피부전도도는 250 Hz로 측정된 신호를 20 Hz로 다운샘플링(Down-sampling)하여 특징을 추출하였다. 감정에 따른 전압 차이를 보기 위해 피부전도도의 평균 및 표준편차를 특징으로 선택하였으며 이는 Equation 9와 Equation 10에 의하여 구할 수 있다. 아래 식들의 N 은 샘플 수를 의미한다. 피부전도도의 영점교차율(Zero-crossing Rate)은 ZC 로 나타내었으며 신호 파형이 영점 축과 교차하는 횟수를 의미하는 특징이다. 영점교차율은 감정의 변화에 따라 교감신경이 활발해지며 진동이 일어나고, 이를 통해 에너지가 집중되는 주파수 영역을 보여준다. 영점교차율은 신호의 미분(Differential)을 통하여 검출할 수 있고 이는 Equation 11를 이용하여 구한다. s 는 신호의 크기, w 는 창함수이며 N 개의 샘플된 신호에서 n 부터 m 까지의 구간의 영점교차율을 구한다(Lisetti & Nasoz, 2004). 피부전도도에서 추출된 특징은 평균, 표준편차, 영점교차율로 총 3개이다.

$$MEAN_{GSR} = \frac{1}{N} \sum_{n=1}^N GSR_n \quad (9)$$

$$SD_{GSR} = \sqrt{\frac{1}{N-1} \sum_{n=1}^N (GSR_n - MEAN_{GSR})^2} \quad (10)$$

$$ZC_{GSR} = \frac{1}{N} \sum_{n=m-N+1}^m \frac{|GSR\{s(n)\} - GSR\{s(n-1)\}|}{2} w(m-n) \quad (11)$$

$$GSR\{s(n)\} = \begin{cases} +1, & s(n) \geq 0 \\ -1, & s(n) < 0 \end{cases}$$

피부 온도는 감정에 따라 체온이 올라가거나 떨어지는 반응을 보이기 때문에 구간에 따른 온도의 평균과 표준편차 총 2가지를 특징으로 선정했다. 피부 온도의 특징은 Equation 12와 Equation 13에 의하여 구할 수 있으며 T 는 시간을 나타낸다(Haag et al., 2004).

$$MEAN_{SKT} = \frac{1}{T} \sum_{t=1}^T SKT_t \quad (12)$$

$$SD_{SKT} = \sqrt{\frac{1}{T-1} \sum_{t=1}^T (SKT_t - MEAN_{SKT})^2} \quad (13)$$

추출된 특징 값들의 범위는 다양하기 때문에 특정 생체신호에 영향을 받지 않도록 정규화 과정이 필요하

다(Haag et al., 2004). 24가지 특징을 가진 400개의 데이터는 Equation 14에 의하여 정규화된다. Equation 14에서 i 는 특징개수로 총 24개의 특징에 대하여 정규화 과정을 거치며, $D(i)$ 는 원래 데이터의 값, $D_{\min}(i)$ 는 각 특징에서의 최소값, $D_{\max}(i)$ 는 각 특징에서의 최대값, $D_{norm}(i)$ 는 정규화 된 데이터를 나타낸다. 정규화 과정을 거친 특징값들은 0에서 1의 범위를 가지게 된다.

$$D_{norm}(i) = \frac{D(i) - D_{\min}(i)}{D_{\max}(i) - D_{\min}(i)} \quad (14)$$

4가지 감정 별 조건에 따라 추출되는 24개의 특징 값들에 대해 일원분산분석과 Tukey 사후검정을 시행하였다. 일원분산분석 결과 맥파의 SDNN, RMSSD, 주파수파워비(LF/HF), 맥파전달속도의 표준편차, 체온 평균 특징에 대해서 일부 피험자의 경우 유의미한 차이를 확인하지 못하였으나 나머지 특징들에서는 유의미한 차이($P < 0.05$)를 확인하였다. Tukey 사후검정 결과에서도 맥파의 SDNN, RMSSD, 주파수파워비(LF/HF), 맥파전달속도의 표준편차, 체온 평균 특징을 제외한 나머지 특징들에서는 정서(Valence)에 따라서 감정 그룹간의 유의미한 차이($P\text{-value} < 0.05$)를 확인하였다.

통계적인 방법으로 특징에 따라 피험자별 감정분류의 차이를 확인하는 데는 한계가 있어 감정 패턴 분류기를 설계하여 감정을 판별하고자 한다. 분류기의 성능은 정규화 된 데이터를 이용하여 5-fold Cross Validation 과정을 통해 검증하였다.

2.3. 감정 패턴 분류기 구성

DBN은 Deep Learning의 한 종류로 다층 신경회로망과 유사한 구조를 가진다. DBN은 Restricted Boltzmann Machine (RBM)을 다층으로 쌓아 올려 구성한다. RBM은 확률적으로 해석한 에너지 모델로서 하나의 입력층과 하나의 은닉층으로 구성되어 있다. RBM의 입력층은 고정되어있으나 은닉층은 자유로운 상태를 가진다. 따라서 입력층의 고정값을 이용한 확률분포를 따라 비교사학습을 진행하여 은닉층의 값을 추정할 수 있게 된다. Equation 15에서 볼 수 있듯이 예너

지는 낮은 방향으로 이동하므로 확률은 커지는 방향으로 진행되며, RBM 학습 역시 에너지가 낮은 방향으로 정상상태가 될 때까지 학습되어 진다(Hinton et al., 2006). 여기서 w 는 가중치, v 는 입력층 노드 값, h 는 은닉층 노드 값, b 는 입력층 바이어스, c 는 은닉층 바이어스, i 는 입력층 노드 순서, j 는 출력층 노드 순서를 나타낸다(Chen & Murray, 2003).

$$P(v, h) \propto e^{-E(v, h)} \quad (15)$$

$$E(v, h) = - \sum_{i, j} v_i h_j w_{ij} = -b^T v - c^T h - v w h^T \quad (16)$$

DBN은 층을 쌓는 초기 과정에서부터 가중치를 조정하지만 오류 역전파 알고리즘(Back-propagation Algorithm, BP Algorithm)은 최종값을 통하여 오차를 최소화하는 방향으로 가중치를 조정한다는 차이점을 가진다(Hinton et al., 2006). DBN은 다층 신경회로망(Multilayer Neural Networks, MNN)과 유사한 구조를 가진다. MNN은 입력층과 한 개 이상의 은닉층 그리고 출력층으로 구성되어 있으며, 은닉층은 복잡한 문제의 학습을 가능하게 하는 상호연결의 정도를 나타내는 역할을 한다. MNN은 오류 역전파 알고리즘을 사용하여 전달된 출력 값 및 본래의 출력 값을 비교하여 오차를 줄여주는 방향으로 가중치를 갱신시킨다.

DBN은 RBM 층의 학습이 모두 끝난 후 오류 역전파 알고리즘을 사용하여 갱신시키는 변별 미세조정(Discriminative Fine-tuning) 과정을 거치며, 이는 층별 탐욕 훈련(Greedy Layer-wise Training)이다. 층별 탐욕 훈련은 RBM 층을 올려가며 노드 값을 확률적으로 샘플링시키는 것이며 이 과정을 통하여 가중치, 바이어스 값이 계산된다. 이러한 과정을 통하여 결정된 가중치 및 바이어스 값들은 MNN의 초기값으로 설정되며 오류 역전파 알고리즘을 통하여 다시 갱신된다(Hinton et al., 2006).

본 논문에서는 이와 같은 특성을 가지는 DBN을 이용하여 감정 패턴 분류기를 구성했다. RBM은 개인 또는 영상에 의하여 발생하는 감정을 확률적으로 판단하여 가중치를 갱신하며, 입력으로 사용된 특징들을 다른 차원으로 매핑시켜주는 역할을 한다. RBM을

통하여 계산된 가중치는 최종적으로 미세조정(Fine-tuning) 과정을 거쳐 은닉층 노드 값을 설정할 수 있게 된다(Hinton et al., 2006). 이러한 과정을 통하여 비선형적인 감정을 판별할 수 있는 감정 분류기를 설계할 수 있고, 기존 오류 역전파 알고리즘의 지역 최적화 문제점을 극복할 수 있다.

Fig. 1은 DBN 구조를 이용한 감정 패턴 분류기 설계의 전체적인 모식도이며, HRV는 심박변이도, PTT는 맥파 전달시간, HF/LF는 고주파파워와 저주파파워의 비, SD는 표준편차, Zero-crossing은 영점교차율 그리고 BP는 오류 역전파 알고리즘을 나타낸다.

2.4. Deep Belief Network 설계

DBN은 RBM을 쌓아 구성되기 때문에 RBM의 초기조건 설정에 따라 다른 성능을 가진다.

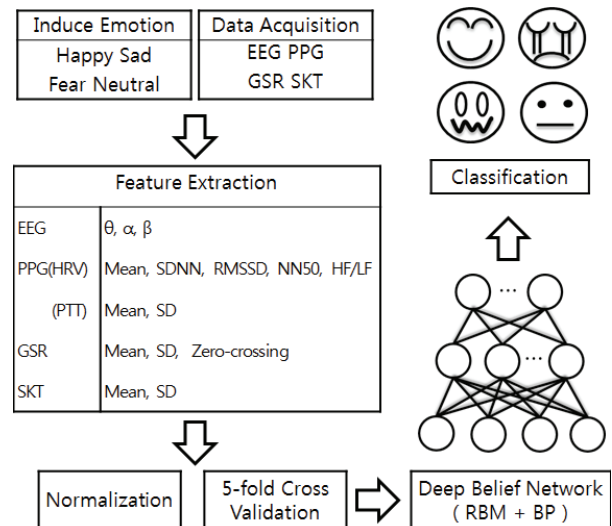


Fig. 1. Design diagram of emotion patterns classifier

먼저 DBN은 RBM의 노드 개수 및 층수에 따라 성능이 결정되는데 노드 개수와 층수가 높을수록 복잡도가 증가하므로 판별의 정확도를 높이지만 너무 높은 층을 올리게 되면 오히려 알고리즘 성능이 떨어지는 결과를 가지게 되므로 적당한 노드 개수를 찾고 RBM 층을 쌓는 것이 중요하다. 두 번째로 학습률은 너무 낮으면 알고리즘 속도가 느려지고 너무 크면 지역 최적점에 빠지는 문제를 가진다. 세 번째로 초기 입력을 데이터 하나하나 입력하는 패턴모드(Pattern

Mode)로 선택할 것인가 한꺼번에 여러 개의 데이터를 입력으로 사용하는 배치모드(Batch Mode)를 선택할 것인가에 따라 성능이 바뀔 수 있다. 적당한 크기의 배치데이터 크기는 알고리즘의 성능을 높이며 학습 시간을 줄일 수 있다. 마지막으로 RBM 반복 횟수 역시 알고리즘 속도와 관련되므로 수렴상태가 되었을 때 멈춰주는 것이 좋으며, 각각의 조건들의 설정을 통하여 성능과 속도가 효율적인 적당한 DBN을 설계할 수 있다(LeCun & Ranzato, 2013).

DBN을 이용한 감정 패턴 분류기를 설계하기 위하여 임의로 RBM의 초기 조건을 설정하여 주었다. 본 논문에서는 학습률은 0.1, 배치데이터의 수는 10개, RBM의 반복 횟수는 10번으로 임의 설정 값을 통하여 노드 개수와 RBM 층의 개수를 먼저 정하였다.

Fig. 2의 왼쪽 상단은 첫 번째 층, 오른쪽 상단은 두 번째 층, 왼쪽 하단은 세 번째 층, 오른쪽 하단은 네 번째 층을 의미한다. 그래프는 노드 개수와 학습 횟수에 따른 오차를 맵을 통하여 표현하였으며 맵의 색이 진할수록 오차가 크고 색이 연할수록 오차가 작다. Fig. 2를 보면 노드 수가 많아질수록 시작 오차가 작아진다. 그러나 열 번의 반복을 거쳤을 경우 노드 수가 커질수록 노드 간 오차가 줄어들는다. 첫 번째 층의 경우에는 시작 오차와 상관없이 열 번의 반복을 거치면 오차가 비슷해지는 것을 볼 수 있으므로 적당한 열 번의 반복을 통한 오차 수렴 정도와 계산 복잡도를 고려하여 첫 번째 층의 노드 개수는 4개로 설정했다.

4개의 노드를 가지는 RBM 층을 쌓은 후 두 번째 층

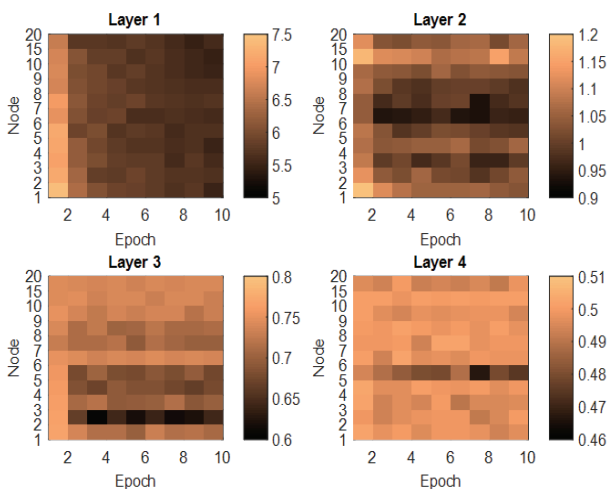


Fig. 2. Error change graphs for the number of nodes by each layer

역시 첫 번째 층과 같이 오차를 구했다. 두 번째 층 역시 노드 개수가 늘어날수록 시작 오차가 줄어들지만, 노드 개수가 어느 정도 이상이 되면 시작 오차가 커지고 열 번 반복 후에도 오차가 다른 경우보다 커지는 것을 확인할 수 있다. 따라서 열 번 반복 후 오차가 가장 작은 3개를 두 번째 층의 노드 개수로 설정했다.

세 번째 층 및 네 번째 층의 오차는 노드 개수가 변화해도 큰 차이를 보이지 않았으며 반복 횟수에 따라서도 많은 변화가 없었다. 따라서 세 번째 층은 열 번의 반복 후 가장 작은 오차를 가지는 2개의 노드를 선택했다.

RBM 구성 시 학습률은 지역 최적화를 막아주고 속도를 조절해주는 역할을 하므로 각 층별로 반복 횟수에 따라 오차를 비교했다. Fig. 3의 왼쪽 상단은 첫 번째 층, 오른쪽 상단은 두 번째 층, 왼쪽 하단은 세 번째 층을 의미하며, 학습률과 학습 횟수에 따른 오차맵을 표현한 것이다.

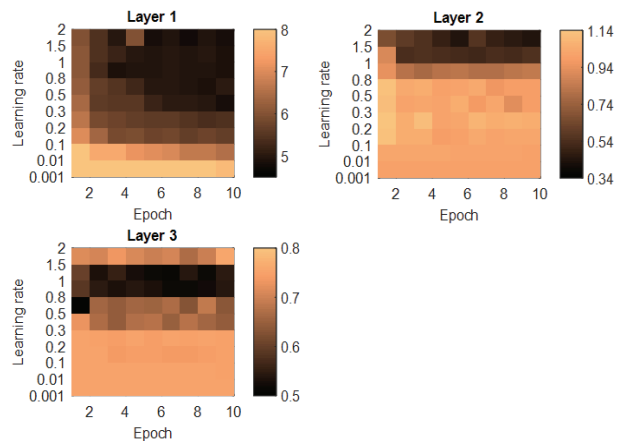


Fig. 3. Error change graphs for learning rate by each layer

첫 번째 층과 두 번째 층에서의 오차를 보면 학습률이 작은 경우에는 반복해도 오차에 큰 변화가 나타나지 않는다. 그러나 학습률이 높은 경우에는 오차가 크게 줄어들을 볼 수 있다. 세 번째 층에서의 결과를 확인하면 학습률이 높은 경우에는 오차의 변화가 있음을 알 수 있다. 하지만 학습률이 1 이상인 경우 오차가 다른 경우에 비해 높은 것을 볼 수 있으며 같은 반복 횟수 내에서 지역 최적화 될 가능성을 가지므로 학습률은 0.5로 설정했다.

다음으로 RBM의 배치데이터의 개수를 결정하기 위하여 데이터를 한 개씩 넣어주는 패턴모드에서 학

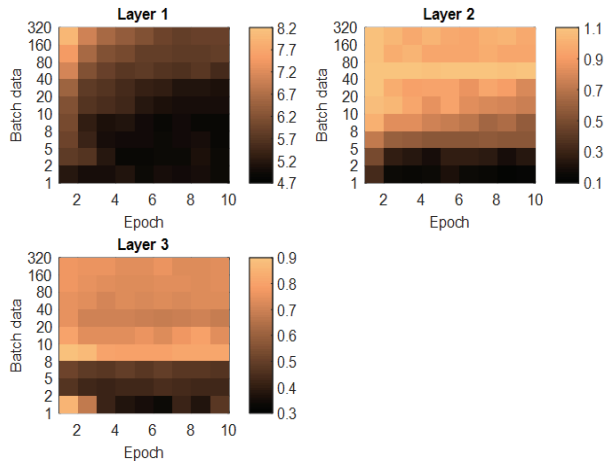


Fig. 4. Error change graphs for the number of batch data by each layer

습 데이터의 개수인 320개를 모두 넣는 경우까지를 비교하여 보았다.

Fig. 4의 왼쪽 상단은 첫 번째 층, 오른쪽 상단은 두 번째 층, 왼쪽 하단은 세 번째 층을 의미하며 배치데이터의 개수와 학습 횟수에 따른 오차맵을 통하여 표현한 것이다. 배치데이터의 개수는 줄어들수록 오차가 작아짐을 세 개의 층 모두에서 확인할 수 있다. 세 번째 층에서는 마지막 반복 시 오차가 배치데이터의 개수에 따라 차이를 가졌으나 시작 오차는 큰 차이를 보이지 않음을 확인할 수 있다. 따라서 배치데이터의 개수는 알고리즘의 속도를 고려하여 오차가 크게 차이가 없는 20개로 설정했다.

RBM의 반복 횟수는 알고리즘의 속도와 관련이 있으므로 오차가 일정한 값으로 수렴하게 된다면 필요하지 않은 반복은 줄여도 된다. 따라서 반복 횟수에 따른 오차 수렴 정도를 비교하였으며 Fig. 5는 반복 횟수에 따른 오차를 층별로 나타내 준 것이다.

첫 번째 층에서는 여덟 번째 반복 이후로 오차가 거의 일정한 값을 가지며 두 번째 층과 세 번째 층에서는 약 세 번째 반복 이후로 일정한 수렴 값 사이에서 오차가 변화하는 것을 볼 수 있다. 이에 따라 첫 번째 층에서의 오차를 고려하여 반복횟수를 10번으로 설정했다.

위와 같은 결과들을 통하여 본 논문에서는 감정 패턴 분류를 위한 DBN을 설계하기 위한 최종 조건으로 총 층의 개수는 3개, 각 층의 노드 개수는 첫 번째 층부터 4개, 3개, 2개로 설정하였으며, 학습률은 0.5, 배치데이터의 개수는 20개, 반복 횟수는 10번으로 설정하였다.

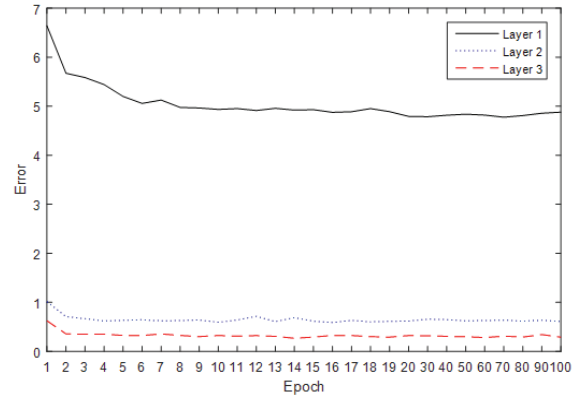


Fig. 5. Error graph depending on the number of repetition

3. 결과

3.1. DBN과 BP 성능 분석

본 논문에서는 DBN의 우수성을 확인하기 위하여 유사한 방법으로 학습되지만 RBM을 사용하지 않는 오류 역전과 알고리즘을 통하여 두 알고리즘의 성능을 분석해보고자 했다. 또한, 시각 자극에서 감정 유발이 잘 되는 구간을 학습데이터로 사용하여 DBN 및 오류 역전과 알고리즘을 학습한 후 같은 5개의 테스트 데이터를 이용하여 성능을 분석했다.

Table 1은 5-fold Cross Validation을 통하여 각 알고리즘의 정확도를 구하고 이의 평균을 구하여 나타낸 것이며, BP는 오류 역전과 알고리즘, P-DBN은 감정 유발이 잘 되는 구간을 학습데이터로 사용한 DBN, P-BP는 감정 유발이 잘 되는 구간을 학습데이터로 사용한 오류 역전과 알고리즘을 나타낸다.

Table 1에서 볼 수 있듯이 DBN은 모두 90% 이상으로 평균 94%의 정확도를 가졌으며 오류 역전과 알고리즘의 평균 정확도인 40% 보다 50% point 더 높은 정확도를 가지는 것을 확인할 수 있다. 감정 유발이 잘 되는 구간을 학습데이터로 사용한 경우 DBN의 결과가 가장 우수함을 확인할 수 있었으며 DBN의 경우보다 평균 정확도가 약 0.75% 더 높은 것을 볼 수 있었지만 뚜렷하게 큰 차이는 보이지 않았다.

3.2. 생체신호에 따른 DBN 결과 분석

본 논문에서는 뇌파, 맥파, 피부전도도, 피부 온도에

Table 1. Algorithm accuracy (%)

Algorithm	Accuracy 1	Accuracy 2	Accuracy 3	Accuracy 4	Accuracy 5	Mean Accuracy
DBN	95.00	90.00	95.00	98.75	91.25	94.00
BP	50.00	25.00	50.00	50.00	25.00	40.00
P-DBN	98.75	92.50	98.75	91.25	92.50	94.75
P-BP	50.00	50.00	25.00	50.00	26.25	40.25

기반을 둔 다양한 생체신호로부터 감정 패턴 분류기의 입력데이터를 추출했다. 어떠한 생체신호가 입력으로 사용되는가에 따라 분류기의 성능은 변화한다. 따라서 생체신호에 따른 DBN의 성능을 분석하고자 하였으며 이는 Table 2를 통하여 나타내었고, 정확도는 5-fold Cross Validation의 평균 정확도를 사용했다.

추출한 모든 생체신호인 뇌파, 맥파, 피부전도도, 피부 온도를 입력으로 사용한 경우 DBN은 약 94%의 가장 높은 정확도를 가졌다. 다음으로 피부 온도를 제외한 뇌파, 맥파, 피부전도도를 사용하였을 경우 약 87.75%의 정확도를 가졌다. 가장 낮은 정확도를 가지는 경우는 피부 온도만 사용한 경우로 약 25%의 정확도를 가졌으며, 맥파만 사용하였을 경우 역시 33%로 낮은 정확도를 가졌다.

한 종류의 생체신호만 사용한 경우 뇌파를 제외하고 모두 50% 미만의 정확도를 가졌으며 뇌파 역시 51%의 정확도를 가짐을 볼 수 있다. 두 종류의 생체신호를 사용한 경우는 한 종류의 생체신호만 사용한

경우와 비교하여 정확도가 높아졌으나 맥파와 피부 온도를 생체신호로 사용한 경우 37.5%의 낮은 정확도를 가지는 것을 확인할 수 있다. 이와 반대로 뇌파와 피부전도도를 사용한 경우 78.5%의 비교적 높은 정확도를 가졌으며 이는 뇌파, 맥파 그리고 피부 온도를 사용한 경우와 같은 정확도를 가짐을 볼 수 있다. 세 종류의 생체신호를 사용한 경우는 모두 오류 역전과 알고리즘을 사용한 결과보다 우수한 정확도를 가짐을 확인할 수 있었으나 뇌파를 제외한 생체신호들을 사용한 경우 57.25%로 뇌파를 생체신호로 선택한 경우에 비하여 낮은 정확도를 보임을 알 수 있다.

4. 고찰 및 결론

본 연구의 목적은 영상정보의 시각자극에 의해 유발된 감정을 생체신호를 통해 확률적으로 해석하여 정확도가 높은 감정 패턴 분류기를 구현하는 것이다.

인간은 다양한 감정을 느끼며 개인의 차가 크므로 분류에 대한 감정 판단의 기준이 모호하기 때문에 정확한 감정 판별이 어렵다. 또한, 시각 영상 자극을 통한 감정 판별은 시간에 따라 영상 자극이 변화하기 때문에 감정 측정 시 계속 실험자의 의도대로 감정이 유도되는지 확인할 수 없다는 한계를 가진다.

본 연구에서는 감정 판별의 한계를 해결하기 위하여 감정 자극 시 중추신경계 및 자율신경계의 생체신호를 종합적으로 획득하여 객관적인 판별을 하고자 했다. 또한, 영상 자극에서 획득된 감정 신호를 연속적으로 보기 위하여 생체신호를 창 이동(Sliding Window) 시켜 주어 특징을 추출했다. 개인 또는 영상에 따른 감정의 모호함을 해결하기 위해 사용한 방법인 비교사학습을 통하여 진행되는 RBM으로 확률을 기반으로 감정을 판별하여 주었다(Chen & Murray, 2003). 이는 각각의 시각 영상 자극에서 영상의 분위기를 통한 해석을 가

Table 2. Classification by bio-signals accuracy (%)

EEG	PPG	GSR	SKT	Accuracy
o	o	o	o	94.00
o	o	o		87.75
o	o		o	78.50
o		o	o	81.75
	o	o	o	57.25
o	o			77.75
o		o		78.50
o			o	52.25
	o	o		52.00
	o		o	37.50
		o	o	57.75
o				51.00
	o			30.00
		o		44.00
			o	25.00

능하게 해주며 시간에 따른 영상의 변화에 관한 한계를 극복할 수 있다. 또한, 피험자 역시 각각의 감정을 느끼는 데 개인적인 차이를 가지고 있는데 감정의 차를 확률적으로 해석함으로써 감정 판별의 한계를 극복했다(Hinton et al., 2006).

먼저 DBN의 설계를 위해서는 여러 조건이 필요하다. 입력 데이터에 따라 DBN의 노드와 층의 복잡도가 결정되며, 결정된 DBN의 구조에서 더욱 우수한 성능을 가지기 위하여서는 학습률, 배치데이터의 개수, 학습 횟수 등을 조절해 줄 수 있다. DBN은 RBM을 기본으로 사용하기 때문에 RBM의 필요조건들을 다양하게 설정해 준 후, 반복 수행을 통하여 RBM을 쌓아 올림으로 성능이 우수한 DBN을 설계할 수 있다(Hinton et al., 2006).

본 논문에서는 반복 수행을 통한 DBN의 설계 결과 평균 정확도가 94%인 감정 패턴 분류기의 구현이 가능했다. 또한, 이는 RBM을 사용하지 않고 랜덤한 가중치 값을 통하여 분류하는 오류 역전파 알고리즘의 평균 정확도인 40% 보다 50% point 높은 정확도를 가지므로 감정 패턴 분류기로서 우수한 성능을 가짐을 알 수 있다.

생체신호 별로 DBN의 성능을 비교한 결과 중추신경계 및 자율신경계의 특징을 모두 사용한 경우가 가장 좋은 성능을 보임을 확인했다. 또한, 생체신호가 하나만 특징으로 들어간 경우 50% 정도 혹은 그 이하의 정확도를 가지며 감정 패턴 분류기로서 좋은 성능을 가지지 못함을 확인하였다. 그러나 중추신경계 신호인 뇌파의 경우 특징으로 뇌파신호만 사용되어도 51%의 정확도를 확인하였으며, 이에 따라 뇌파가 감정 판별을 하는데 주요한 역할을 함을 볼 수 있다. 감정 판별 시 80%가 넘는 정확도를 가지는 모델을 설계하기 위해서는 생체신호를 세 가지 이상 사용하여야 하며, 이때 뇌파와 피부전도도는 반드시 사용되어야 하는 특징으로 볼 수 있다. 또한, 뇌파와 피부전도도만을 사용한 경우는 뇌파, 맥파, 피부 온도를 사용한 경우와 같은 정확도를 보이므로 실험의 복잡도 및 계산 복잡도를 고려한다면 분류기가 같은 성능을 내는 경우 뇌파와 피부전도도를 통한 감정의 분류도 가능할 것으로 사료된다. 피부 온도는 가장 낮은 정확도를 보였으며 피부 온도를 사용한 경우 정확도는

많이 증가하지 않는 것을 보았다. 이는 피부 온도가 외부 요인에 민감하므로 감정 자극에 의하여서도 변하지만, 외부온도 즉, 실내온도 등의 영향도 받기 때문으로 사료된다.

향후 감정 패턴 분류기의 설계를 위한 실험 시 시각 영상 자극이 모든 피험자에게 작용이 되는지에 관한 고찰이 필요하며 더 효율적인 실험 프로토콜을 통한 데이터 획득이 필요하다. 또한, 데이터 획득 과정에서 피험자는 노년층 여성을 대상으로 하였는데 이는 앞으로 다양한 연령층과 성별을 대상으로 데이터를 획득하여 비교할 수 있을 것으로 보인다.

두 번째로 RBM을 통하여 DBN을 구성하므로 오류 역전파 알고리즘이 가지는 문제점을 해결할 수 있었으나 DBN을 설계하는 과정에서 학습률, 배치데이터 개수, 학습 횟수 등의 가변적인 요소가 존재하였다. 이러한 요소들을 더욱 최적화시켜줄 수 있다면 감정 패턴 분류기의 성능이 개선될 가능성을 가진다. 따라서 이와 같은 가변적인 요소를 최적화시켜줄 수 있는 추가 연구가 필요할 것으로 사료된다.

마지막으로 본 연구에서는 알고리즘을 확률적으로 해석하였으나 한 명의 개인 데이터로부터 확률적 학습을 하였기 때문에 각각의 개인마다 알고리즘의 조건을 최적화해야 한다는 한계점을 가진다. 따라서 추후에는 알고리즘의 일반화를 통하여 모든 피험자 데이터를 만족하게 하는 알고리즘 구현에 관한 연구가 필요하다.

우수한 성능을 가지는 감정 패턴 분류기를 설계하는 방법으로 본 연구에서는 DBN을 이용한 다층 신경회로망을 구성했다. 본 연구에서 설계된 감정 패턴 분류기는 시각 영상 자극을 통한 인간의 감정을 판별할 때 우수한 성능을 가질 것으로 보이며, 추후 인지 과학 및 HCI 분야 등에서 활용 가능할 것으로 보인다(Cowie et al., 2001).

REFERENCES

- Chen, H., & Murray, A. F. (2003). Continuous Restricted Boltzmann Machine with an Implementable Training Algorithm. *In Vision, Image and Signal Processing, IEEE Proceeding-*, 150(3), 153-158.
- Choi, W. (2011). *A Classification Analysis of Negative Emotion Based on PPG Signal Using Fuzzy-Ga.*

- master's thesis, Yonsei University, Seoul.
- Cowie, R., Douglas-Cowie, E., Tsapatsoulis, N., Votsis, G., Kollias, S., Fellenz, W., & Taylor, J. G. (2001). Emotion Recognition in Human-Computer Interaction. *Signal Processing Magazine, IEEE*, 18(1), 32-80.
- Guang-yuan, L., & Min, H. (2009). Emotion Recognition of Physiological Signals Based on Adaptive Hierarchical Genetic Algorithm. In *2009 World Congress on Computer Science and Information Engineering*, 670-674.
- Haag, A., Goronzy, S., Schaich, P., & Williams, J. (2004). Emotion Recognition Using Bio-Sensors: First Steps Towards an Automatic System. In *Tutorial and Research Workshop on Affective Dialogue Systems*, 36-48.
- Hinton, G. E., Osindero, S., & Teh, Y.-W. (2006). A Fast Learning Algorithm for Deep Belief Nets. *Neural Computation*, 18(7), 1527-1554.
- Jerritta, S., Murugappan, M., Nagarajan, R., & Wan, K. (2011). Physiological Signals Based Human Emotion Recognition: A Review. In *Signal Processing and its Applications (CSPA), 2011 IEEE 7th International Colloquium on*, 410-415.
- Khashman, A. (2008). A Modified Backpropagation Learning Algorithm with Added Emotional Coefficients. *Neural Networks, IEEE Transactions on*, 19(11), 1896-1909.
- Kleinginna, P. R., & Kleinginna, A. M. (1985). Cognition and affect: A reply to Lazarus and Zajonc. *American Psychologist*, 40(4), 470-471.
- Krause, R. (1987). Universals and Cultural Differences in the Judgments of Facial Expressions of Emotion. *Journal of Personality and Social Psychology*, 53(4), 712-717.
- Lang, P. J. (1995). The Emotion Probe: Studies of Motivation and Attention. *American Psychologist*, 50(5), 372-385.
- LeCun, Y., & Ranzato, M. (2013). Deep Learning Tutorial. In *Tutorials in International Conference on Machine Learning (ICML13)*, Citeseer.
- Lisetti, C. L., & Nasoz, F. (2004). Using Noninvasive Wearable Computers to Recognize Human Emotions from Physiological Signals. *EURASIP Journal on Advances in Signal Processing*, 11, 1-16.
- Malik, M., & Camm, A. J. (1990). Heart Rate Variability. *Clinical Cardiology*, 13(8), 570-576.
- Moretti, D. V., Babiloni, C., Binetti, G., Cassetta, E., Dal Forno, G., Ferreric, F., & Nobili, F. (2004). Individual Analysis of Eeg Frequency and Band Power in Mild Alzheimer's Disease. *Clinical Neurophysiology*, 115(2), 299-308.
- Murugappan, M., Ramachandran, N., & Sazali, Y. (2010). Classification of Human Emotion from Eeg Using Discrete Wavelet Transform. *Journal of Biomedical Science and Engineering*, 3, 390-396.
- Niu, X., Chen, L., & Chen, Q. (2011). Research on Genetic Algorithm Based on Emotion Recognition Using Physiological Signals. In *2011 International Conference on Computational Problem-Solving*, 614-618.
- Peng, Y., Zhu, J.-Y., Zheng, W.-L., & Lu, B.-L. (2014). Eeg-Based Emotion Recognition with Manifold Regularized Extreme Learning Machine. In *Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE*, 974-977.
- Schaaff, K., & Schultz, T. (2009). Towards Emotion Recognition from Electroencephalographic Signals. In *2009 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops*, 1-6.
- Wang, D., & Shang, Y. (2013). Modeling Physiological Data with Deep Belief Networks. *International of Journal Information and Education Technology (IJJET)*, 3(5), 505-511.

원고접수: 2016.03.14

수정접수: 2016.12.08

게재확정: 2017.03.21