

CNN 기반 인간 동작 인식을 위한 생체신호 데이터의 증강 기법

게렐바트¹, 권춘기^{1*}

¹순천향대학교 의료IT공학과

Bio-signal Data Augmentation Technique for CNN based Human Activity Recognition

Gerelbat BatGerel¹, Chun-Ki Kwon^{1*}

¹Department of Medical IT Engineering, Soonchunhyang University

요약 합성곱 신경망을 비롯하여 딥러닝 신경망의 학습에서 많은 양의 훈련데이터의 확보는 과적합 현상을 피하고 우수한 성능을 가지기 위해서 매우 중요하다. 하지만, 딥러닝 신경망에서의 레이블화된 훈련데이터의 확보는 실제로는 매우 제한적이다. 이를 극복하기 위해, 이미 획득한 훈련데이터를 변형, 조작 등으로 추가로 훈련데이터를 생성하는 여러 증강 방법이 제안되었다. 하지만, 이미지, 문자 등의 훈련데이터와 달리, 인간 동작 인식을 행하는 합성곱 신경망의 생체신호 훈련데이터를 추가로 생성하는 증강 방법은 연구 문헌에서 찾아보기 어렵다. 본 연구에서는 합성곱 신경망에 기반한 인간 동작 인식을 위한 생체신호 훈련데이터를 생성하는 간편하지만, 효과적인 증강 방법을 제안한다. 본 연구의 제안된 증강 방법의 유용성은 추가로 생성된 생체신호 훈련데이터로 학습하여 합성곱 신경망이 인간 동작을 높은 정확도로 인식하는 것을 보임으로써 검증하였다.

• 주제어 : 생체신호, 데이터 증강, 합성곱 신경망, 인간 동작 인식, 딥러닝

Abstract Securing large amounts of training data in deep learning neural networks, including convolutional neural networks, is of importance for avoiding overfitting phenomenon or for the excellent performance. However, securing labeled training data in deep learning neural networks is very limited in reality. To overcome this, several augmentation methods have been proposed in the literature to generate an additional large amount of training data through transformation or manipulation of the already acquired training data. However, unlike training data such as images and texts, it is barely to find an augmentation method in the literature that additionally generates bio-signal training data for convolutional neural network based human activity recognition. Thus, this study proposes a simple but effective augmentation method of bio-signal training data for convolutional neural network based human activity recognition. The usefulness of the proposed augmentation method is validated by showing that human activity is recognized with high accuracy by convolutional neural network trained with its augmented bio-signal training data.

• Key Words : Bio-signal, Data augmentation, Convolutional neural network, Human activity recognition, Deep learning

Received 19 June 2023, Revised 26 June 2023, Accepted 30 June 2023

* Corresponding Author Chun-Ki. Kwon, Department of Medical IT Engineering, Soonchunhyang University, 22, Soonchunhyang-ro, Asan-si, Chungcheongnam-do, Korea. E-mail: chunkikwon@sch.ac.kr

I. 서론

합성곱 신경망을 비롯한 딥러닝 신경망은 이미지 인식을 포함해서 다양한 분야에서 우수한 성능을 보여왔다[1-3]. 이러한 딥러닝 신경망의 우수한 성능은 많은 양의 훈련데이터가 필요하다. 하지만, 현실적으로 대부분의 응용에서 훈련데이터의 확보는 제한적이며, 이러한 어려움을 극복하기 위하여 기존 대부분의 증강 기법은 확보된 데이터들을 변형, 노이즈 신호 추가를 하는 등의 복잡하고 개발자의 경험이나 시행착오 등에 많이 의존한다[4-6]. 따라서, 이미 획득한 훈련데이터로부터 손쉽게 데이터를 생성하는 간편하면서도 유용한 증강 기법이 필요하다.

한편, 생체신호인 표면 근전도 신호를 획득하여 인간 동작인 한국숫자지화 제스처를 합성곱 신경망을 활용하여 인식하는 연구를 수행한 참고문헌[7]에서는 표면 근전도 신호를 획득하기 위하여 샘플링 주파수의 적정성을 살펴보았다. 거기서, 측정 장비의 최대 샘플링 주파수 1,024Hz로 획득한 표면 근전도로 학습한 합성곱 신경망의 한국숫자지화 제스처의 인식 성능을 512Hz, 256Hz, 128Hz, 64Hz의 4가지 샘플링 주파수로 획득한 표면 근전도로 학습한 합성곱 신경망의 인식 성능과 비교하였을 때, 학습속도의 차이가 다소 있었지만 한국숫자지화 제스처의 인식 성능은 차이가 없음을 보였다. 이러한 연구 결과는 최대 샘플링 주파수로 획득한 생체신호는 최대 샘플링 주파수보다 낮은 샘플링 주파수에 해당하는 생체신호의 다수의 데이터로 생성할 수 있음을 말한다.

본 연구에서 제안하는 생체신호 훈련데이터에 대한 증강 기법은 선행연구인 참고문헌[7]의 연구결과에서 구체화한 것으로, 합성곱 신경망을 활용하여 인간 동작을 인식하기 위한 생체신호의 훈련데이터를 추가적으로 생성하는 간단하면서도 효과적인 훈련데이터 증강 기법이다[8].

본 연구의 구성은 다음과 같다. 제2장에서는 본 연구에서 제안하는 증강 기법을 기술하고, 제3장에서는 제안된 증강 기법의 유용성을 검증하였으며, 마지막으로 제4장에서는 결론과 향후 연구 방향을 서술하였다.

II. 제안하는 생체신호 훈련데이터 증강 기법

본 연구에서 제안하는 생체신호 훈련데이터의 증강 기법의 기본 개념은 다음과 같다. 생체신호를 측정 장비의 최대 샘플링 주파수로 측정하고 훈련데이터로 생성한 후에 최대 샘플링 주파수보다 낮은 샘플링 주파수로 측정된 것처럼 훈련데이터를 재배열하면 최대 샘플링 주파수와 최대 샘플링 주파수보다 낮은 주파수 간의 비율만큼 생체신호 훈련데이터가 생성되는 것으로서, 최대 샘플링 주파수보다 낮은 주파수로 생성된 훈련데이터가 합성곱 신경망의 훈련데이터로써 활용될 수 있음은 이미 검증되었다[7].

인간 동작을 수행하면 측정 장비의 샘플링 주파수로 여러 생체신호원으로부터 측정된 생체신호는 전처리 과정을 거친 후에 이미지화 과정을 통해 합성곱 신경망의 학습을 위해 이미지화된 생체신호 훈련데이터로 만들어진다. 생체신호 훈련데이터의 크기는 생체신호원의 개수와 샘플링 주파수에 의해 결정된다.

본 연구에서 제안하는 생체신호 훈련데이터의 증강 기법을 설명하기 위하여, 인간 동작을 수행할 시의 생체신호는 m 개의 생체신호원으로부터 측정 장비의 최대 샘플링 주파수를 n Hz로 측정하였다고 가정하면, 그림 1과 같이 $m \times n$ 크기의 생체신호 훈련데이터로 변환된다.

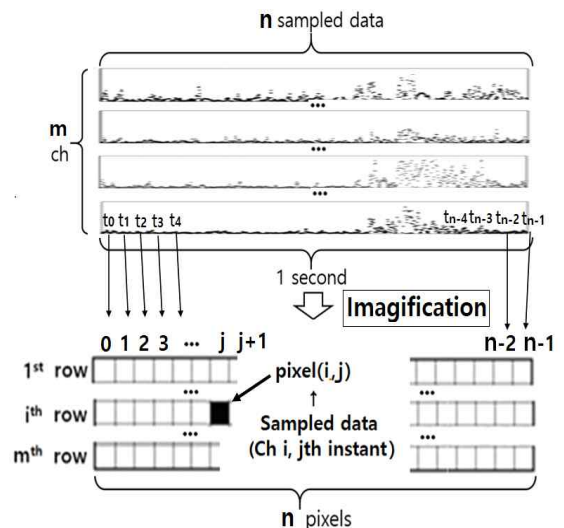


Fig. 1. n Sampled bio-signal data measured for a second from m channels

최대 샘플링 주파수로 측정된 생체신호 훈련데이터로부터 추출하고자 하는 생체신호 훈련데이터의 샘플링 주파수를 $p\text{Hz}$ 라고 하면, 최대 샘플링 주파수 n 은 p 의 정수배가 되며, 이 정수배를 k 라고 표시한다. 따라서, 최대 샘플링 주파수와 최대 샘플링 주파수보다 낮은 샘플링 주파수와의 관계는 식 (1)이다.

$$n = k \cdot p \quad (1)$$

따라서, 최대 샘플링 주파수 $n\text{Hz}$ 로 측정된 하나의 생체신호 훈련데이터의 크기는 $m \times (k \cdot p)$ 가 된다. 샘플링 주파수 $p\text{Hz}$ 의 샘플링 주파수에 해당하는 생체신호 훈련데이터는 최대 샘플링 주파수 $n\text{Hz}$ 로 측정된 하나의 훈련데이터에서 k -배수번째 열의 위치에 있는 $\{0, k, 2k, 3k, \dots, (p-3)k, (p-2)k, (p-1)k\}, \{1, k+1, 2k+1, 3k+1, \dots, (p-3)k+1, (p-2)k+1, (p-1)k+1\}, \dots, \{k-1, 2k-1, 3k-1, \dots, (p-2)k-1, (p-1)k-1, pk-1\}$ 으로 총 k 쌍의 데이터 셋을 얻게 된다. 따라서, 최대 샘플링 주파수 $n\text{Hz}$ 로는 하나의 훈련데이터만 획득하였지만, 제안된 증강기법으로 샘플링 주파수 $p\text{Hz}$ 로 k 개의 훈련데이터가 생성된다. 이러한 과정을 그림 2에 도식화하였다.

이러한 과정을 최대 샘플링 주파수 $n\text{Hz}$ 로 획득한 모든 생체신호 훈련데이터에 적용한다면, k 배의 훈련데이터를 추가로 얻을 수 있다. 예를 들면, 참고문헌[4]에서는 네 개의 전극으로부터 1,024Hz로 측정을 한 표면 근전도 신호의 훈련데이터로부터 각 샘플링 주파수로부터 표 1과 같은 훈련데이터가 생성된다.

Table 1. Training data augmented at sampling frequency of 512Hz, 256Hz, 128Hz, and 64Hz compared to 1,024Hz

Sampling frequency, p	training data augmented by		
	generation rate, k	total data	size, $m \times p$
1,024Hz	1,024 / 1,024	200	4x1,024
512Hz	1,024 / 512	400	4x512
256Hz	1,024 / 256	800	4x256
128Hz	1,024 / 128	1,600	4x128
64Hz	1,024 / 64	3,200	4x64

III. 제안 증강 기법의 유용성 검증

본 연구에서 제안하는 생체신호 훈련데이터의 증강 기법의 유용성을 보이기 위하여, 표면 근전도 신호를

활용한 합성곱 신경망에 기반한 한국숫자지화 제스처 인식 시스템을 활용하였으며[7], 최대 샘플링 주파수 1,024Hz로 획득한 생체신호 훈련데이터와 이들 훈련데이터로부터 제안된 증강 기법을 활용하여 생성된 샘플링 주파수 256Hz에 해당하는 훈련데이터로 합성곱 신경망을 학습시키어 합성곱 신경망이 한국숫자지화 제스처를 인식하는 성능을 비교하여 제안된 증강 기법의 유용성을 검증한다.

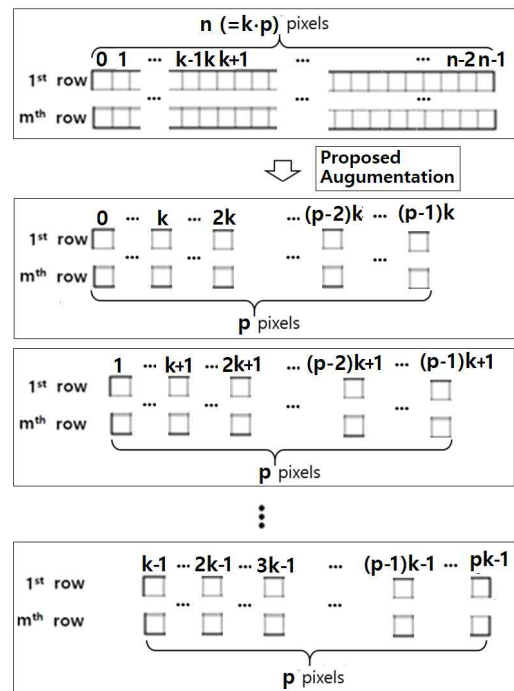


Fig. 2. Proposed augmentation technique of bio-signal training data for CNN based human activity recognition

3.1 인식할 인간동작 - 한국숫자지화 제스처

본 연구에서 채택한 인간 동작은 숫자 일(1)부터 숫자 다섯(5)까지의 다섯 한국숫자지화 제스처들이다. 이들 제스처를 Fig. 3에 나타내었다.

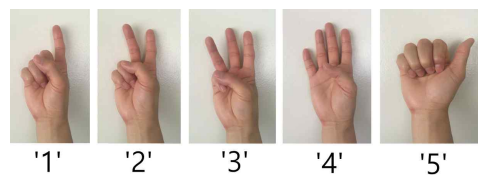


Fig. 3. Five Korean finger number gestures from one to five

3.2 생체신호 - 표면 근전도

본 연구에서는 제안한 증강 기법이 CNN의 훈련데이터가 무엇인지와 관계없이 적용 가능하기 때문에, 선행 연구에서 측정된 표면 근전도 신호와 이를 전처리와 이미지화 과정을 거친 표면 근전도 훈련데이터를 그대로 활용하였다[7][10]. 원 표면 근전도 데이터는 전완근의 얇은손가락굽힘근(flexor digitorum superficialis), 긴엄지굽힘근(flexor pollicis longus), 긴엄지벌림근(abductor pollicis longus), 그리고 손가락펴근(extensor digitorum)상의 피부위에 각각 채널1 전극쌍, 채널2 전극쌍, 채널3 전극쌍, 그리고 채널4 전극쌍을 통해서 측정되었다[10].

3.3 Convolutional Neural Network (CNN) 기반 인간 동작 인식 시스템

본 연구에서 채택한 인간 동작 인식 시스템은 참고 문헌[4]의 표면 근전도 신호를 합성곱 신경망이 학습하여 한국숫자지화 제스처를 인식하는 시스템이다. 이 시스템은 한국숫자지화 제스처를 시연할 때에 관련되는 근육상에 부착된 네 쌍의 전극에서 표면 근전도 신호를 1,024Hz로 획득하고, 전처리 과정과 이미지화 과정을 거쳐 합성곱 신경망의 학습을 위한 훈련데이터로 변환하고, 필요시에 본 연구에서 제안한 증강 기법으로 생성된 훈련데이터를 학습하여 한국숫자지화 제스처를 최종적으로 인식한다. 전체적인 CNN기반한 인간 동작 인식 시스템의 블록도를 그림 4에 도시하였다.

합성곱 신경망은 표면 근전도 신호로부터 특징을 추출하는 합성곱층과 풀링층이 결합된 두 개의 숨긴 층, 하나의 합성곱층, 그로부터 시연된 한국숫자지화를 분류하는 두 개의 완전연결층으로 구성되어 있고, 모든 노드에서 활성화 함수는 ReLU 함수, 풀링층은 평균 함수, 그리고 한국숫자지화의 분류를 위한 마지막 완전연결층에서는 소프트맥스 함수를 적용하였다. 합성곱층의 필터의 사이즈는 4x4, 합성곱층의 필터의 개수는 32, 16, 4개로 설정하였고, 두 개의 완전연결층의 노드 개수는 모두 256개로 설정하였다. 증강 기법을 포함한 한국숫자지화 기반 인식 시스템의 전체적인 도식이 Fig. 4에 나타나 있다[7].

3.4 환경 설정

본 연구에서 합성곱 신경망의 학습은 클라우드 기반의 무료 슈퍼노트북 개발환경과 유사한 구글 Colaboratory 딥러닝 플랫폼을 활용하여 학습을 하였으며, GPU와 TPU는 사용하지 않았다[9]. 합성곱 신경망 학습을 위한 주요 하이퍼파라미터는 다음과 같다.

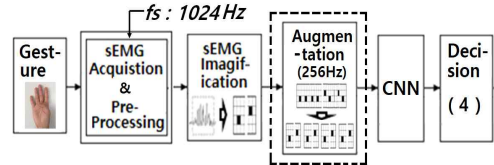


Fig. 4. Block diagram of CNN-based finger number gesture recognition using sEMG signals

합성곱 신경망의 지도학습에서 훈련데이터 셋을 이용하여 합성곱 신경망이 학습할 때, 훈련데이터의 실제 결과와 합성곱 신경망 학습 모델이 예측한 결과에서 발생하는 손실을 계산하는 손실함수는 희박 다분류 교차 엔트로피(Sparse categorical cross-entropy) 이다. 또한 이 함수가 계산한 손실 값을 줄이도록 하는 최적화 함수로는 아담(Adam)을 채택하고, 0.0005의 학습률로 학습 속도를 가시도록 하였습니다. 본 연구에서는 CNN 학습을 에폭 500까지 수행 하였다.

3.5 연구 결과 및 고찰

본 절에서는 증강 기법을 적용하기 전의 원 훈련데이터인 최대 샘플링 주파수 1,024Hz로 획득한 200개의 생체신호 훈련데이터로 한국숫자지화 제스처를 인식하는 합성곱 신경망의 학습 성능은 그림 5에 도시하였다.

그림 6은 원 1,024Hz로 획득한 훈련데이터에 증강 기법을 적용하여 샘플링 주파수 256Hz로 생성·증강된 네 세트 중 하나의 증강 훈련데이터 셋으로 학습한 합성곱 신경망의 학습 성능을 나타내었다. 이것은 네 세트를 개별 세트로 학습한 합성곱 신경망 모델의 경향이 유사하기 때문이다.

그림 7은 원 1,024Hz로 획득한 훈련데이터에 증강 기법을 적용하여 샘플링 주파수 256Hz로 생성·증강된 네 세트를 모두 합친 800개의 증강 훈련데이터로 학습한 합성곱 신경망의 학습 성능을 나타내었다.

그림 5, 그림 6, 그림 7에 도시한 합성곱 신경망의 학습 성능은 지도학습의 매 에폭마다 학습한 합성곱

신경망 모델이 각각의 훈련데이터가 어느 분류에 속하는지를 예측(추론)한 예측 분류 값과 각각의 훈련데이터가 속한 실제 분류 값이 같은 훈련데이터의 개수를 지도학습에 활용한 모든 훈련데이터 개수로 나눈 값인 정확도이다.



Fig. 5 CNN Performance trained with original training data on Korea finger number gesture recognition

그림 5에서 도시하였듯이, 원 1,024Hz로 측정된 표면 근전도로 학습한 한국숫자지화 체크쳐를 분류하는 합성곱 신경망은 에폭 206에서 한국숫자지화 체크쳐를 모두 분류하는 정확도 100%의 합성곱 신경망 모델이 생성되었고 유지되었다.

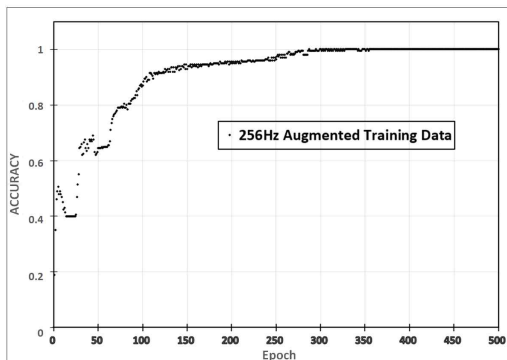


Fig. 6 CNN Performance trained with one of four sets of augmented training data on Korea finger number gesture recognition

그림 6에서, 제안된 증강 기법을 이용하여 생성한 증강 훈련데이터 셋으로 학습한 합성곱 신경망도 에폭 287에서 100% 정확도를 가지는 합성곱 신경망 모델이 생성되었으나, 이후 정확도가 99%대로 유지하면서 에폭 355에서 100% 정확도를 가지는 모델이 만들어진 후 그대로 유지되었다.

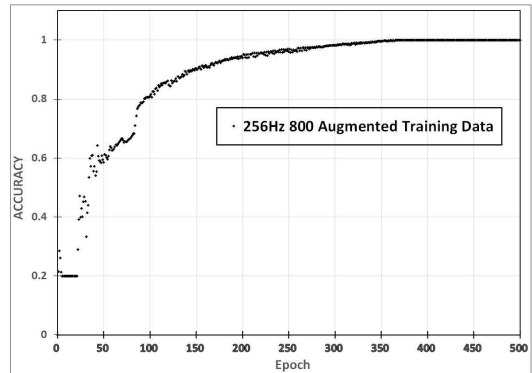


Fig. 7 CNN Performance trained with four sets of augmented training data on Korea finger number gesture recognition

마지막 사례인 제안된 증강 기법을 이용하여 생성한 모든 증강 훈련데이터로 학습한 합성곱 신경망은 훈련데이터의 개수가 다른 경우보다 4배로 많아 상대적으로 느린 학습 속도를 보여주었지만, 에폭 365에서부터 100% 정확도를 가지는 합성곱 신경망 모델이 생성되는 것을 그림 7에서 확인된다.

세 가지 경우 모두, 에폭 50 이전에서 훈련데이터에 대한 학습이 제대로 이루어져서 정확도가 급상승하는 경향을 보였다.

이와 같이, 원 1,024Hz의 학습데이터로 학습한 합성곱 신경망 모델이 학습하는 경향을 제안된 증강기법을 활용하여 생성된 증강 훈련데이터로 학습한 합성곱 신경망 모델에서도 그대로 보여지고 있는 등 학습 속도의 차이가 있지만 합성곱 신경망의 학습 경향이 유사하다.

그림 5, 그림 6, 그리고 그림 7에서의 연구 결과로부터, 본 연구에서 제안한 증강기법에서 생성된 훈련데이터가 원 훈련데이터가 가지고 있는 분류 대상의 정보를 그대로 가지고 있고 합성곱 신경망을 학습할 수 있음을 확인하였다.

따라서, 본 연구에서 제안한 생체신호 훈련데이터에 대한 증강 기법이 간단하면서도 효과적으로 원 훈련데이터로부터 많은 수의 훈련데이터를 생성하여 증강할 필요가 있을 때 유용한 솔루션임을 알 수 있다.

IV. 결론

본 연구에서는 인간 동작 인식 응용에서 합성곱 신경망의 학습에 필요한 생체신호 훈련데이터를 추가로 생성하는 간단하지만, 효과적인 증강 기법을 제안한다. 이 증강 기법은 생체신호를 측정하는 측정 장비가 가용한 최대 샘플링 주파수로 생체신호를 측정하고, 합성곱 신경망을 학습하는 생체신호 훈련데이터는 실질적으로 최대 샘플링 주파수보다 낮은 샘플링 주파수로 샘플링한 것처럼 훈련데이터를 구성한다. 제안된 증강 기법으로 생성된 생체신호 훈련데이터들로 학습을 수행한 합성곱 신경망이 인간 동작을 100%의 정확도로 인식한 연구 결과를 통해 제안된 증강 기법의 유용성을 검증하였다.

기존 증강 기법들이 획득한 원 훈련데이터를 회전, 노이즈 추가 등의 변형·조작 등을 통해서 가상의 훈련데이터를 생성하는 반면, 제안된 증강 기법은 실제로 측정된 데이터를 활용하여 추가로 생성하는 것으로 가상의 훈련데이터가 아닌 실제로 측정된 훈련데이터인 것으로 연구자의 경험이나 시행착오에 의존하지 않는 것이 특징이다. 따라서, 생체신호 훈련데이터를 증강하고자 하는 연구자들에게 기존 증강 기법과 함께 선택할 수 있는 해결책을 제공한다.

ACKNOWLEDGMENTS

본 연구는 부분적으로 2021년도 정부(교육부)의 재원으로 한국연구재단의 지역대학우수과학자지원 사업(No.2021R111A3043994)의 지원을 받아 수행된 것임.

REFERENCES

[1] K. C. Hong, H. S. Kim, and Y. H. Han. (2021, Dec.). CNN-based Sign Language Translation Program for the Deaf. *Journal of Korea Institute of Convergence Signal Processing*, 22(4), pp. 206-212. DOI:10.23087/jkicsp.2021.22.4.009.

[2] H. Park. (2021, Mar.). A Review of 3D Object Tracking Methods Using Deep Learning. *Journal of Korea Institute of Convergence Signal Processing*, 22(1), pp. 30-37.

[3] J. Park and C. Kwon. (2021, Jan.). Korean Finger

Number Gesture Recognition Based on CNN using Surface Electromyography Signals. *Journal of Electrical Engineering & Technology*, 16, pp. 591-598. DOI:10.1007/s42835-020-00587-3.

[4] S. Yang, W. Xiao, M. Zhang, S. Guo, J. Zhao, F. Shen, (2022, April), Image Data Augmentation for Deep Learning:A Survey, arXiv:2204.08610v1.

[5] C. Shorten. T. Khoshgoftaar, and B. Furht. (2021). Text Data Augmentation for Deep Learning. *Journal of Big Data*, 8(101). <https://doi.org/10.1186/s40537-021-00492-0>.

[6] S. Harada, H. Hayashi, and S. Uchida, (2018). Biosignal Data Augmentation Based on Generative Adversarial Networks, 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 18-21 July 2018.

[7] G. Batgerel and C. Kwon, (2023, March), The Study on Effect of sEMG Sampling Frequency on Learning Performance in CNN based Finger Number Recognition, *Journal of Korea Institute of Convergence Signal Processing*, 24(1), pp. 51-56. <https://doi.org/10.23087/jkicsp.2023.24.1.007>.

[8] G. Batgerel and C. Kwon, "Bio-signal Data Augmentation Method for CNN based Human Activity Recognition," in Proc. of the Summer Conf. 2023 on Convergence Signal Processing Technology, Busan, South Korea, 2023.

[9] <https://colab.research.google.com/notebooks/welcome.jpynb>, Google.com.

[10] J. Park and C. Kwon. (2018, Aug.). Study on forearm muscles and electrode placements for CNN based Korean number gesture recognition using sEMG signals. *Journal of Korea Academia- Industrial Cooperation Society*, 19(8), pp. 260-267.

저자소개

게렐바트 (Gerelbat Batgerel)



2015년 6월 : 휴리대학교(몽골)
전자공학과(공학사)
2020년 7월 : 순천향대학교
의료IT공학과(공학석사)
2022년 9월~현재 : 순천향대학교
의료IT공학과(공학박사과정)
관심분야 : 딥러닝 및 신호 응용,

권 춘 기 (Chun-Ki Kwon)



1994년 2월 : 고려대학교
전기공학과(공학석사)
2005년 6월 : 퍼듀대학교
전기공학과(공학박사)
2008년 3월~현재 : 순천향대학교
교수
관심분야 : 딥러닝 및 신호 응용,