

뇌졸중 검출을 위한 복합 생체 정보를 활용한 멀티 모델 프레임워크

Multi-Model Framework based on Complex Biometric Information for Stroke Detection

최형선¹, 김재승², 황보택근^{3*}

HyoungSun Choi¹, JaeSeoung Kim², Taeg-Keun Whangbo^{3*}

요 약

세계보건기구에 따르면 세계 인구는 빠르게 고령화를 향하고 있다. 이는 곧 의료비용의 증가와 각종 만성 질환의 근원지로 예상된다. 세계보건기구에 따르면 뇌졸중은 전 세계 사망원인의 2위를 차지하고 있으며 그 수는 계속해서 증가하고 있다. 이에 각국의 연구원들은 다양한 연구와 임상 실험을 통해 여러 위험 요소가 보고하였고, 뇌졸중의 심각성을 인지하고 있다. 기존의 연구들은 뇌졸중의 증상을 탐지하고, 인과관계를 조사하였다. 또한 인공지능의 발전으로 증상 중 하나인 얼굴의 일그러진 정도를 측정하며 증상 여부를 판별하는데 성공하였다. 그러나 뇌졸중은 얼굴의 마비 증상 외에도 목 소리의 떨림 등 다른 증상들을 될 수 있다. 따라서 본 연구에서는 뇌졸중의 증상 중 말의 어눌함과 얼굴의 일그러진 정도에 초점을 맞추어 인공지능을 활용한 딥러닝 모델 프레임워크를 제안한다. 제안된 모델은 정확도 향상과 부족한 데이터셋의 한계점을 극복하기 위해 전이학습을 적용하였다. 그 결과 훈련 정확도에서 0.7%, 검증 정확도에서는 뇌졸중 환자는 13.9% 일반 환자는 4.6% 개선된 성능을 보였다.

핵심어 : 뇌졸중, 멜-스펙트로그램, 전이학습, 특성추출

Abstract

According to the World Health Organization, the world's population is rapidly aging. This is expected to increase medical costs and be the source of various chronic diseases. According to the World Health Organization, stroke ranks second in the world's death toll and the number continues to rise. Researchers from each country have reported various risk factors through various studies and clinical trials, and are aware of the seriousness of stroke. Previous studies have detected symptoms of stroke and investigated

1 Department Computer Science, Gachon University, Gyeong-Gi, Korea [Graduate Student]
e-mail: hschoi@gachon.ac.kr

2 Health IT Research center, Gachon University Gil Medical Center, Incheon, Republic of Korea [Researcher]
e-mail: mpmkjs@gilhospital.com

3 Department Computer Science, Gachon University, Gyeong-Gi, Korea [Professor]
e-mail: tkwhangbo@gachon.ac.kr (Corresponding author)

* This research was supported by the Bio & Medical Technology Development Program of the National Research Foundation (NRF) funded by the Ministry of Science & ICT(2017M3A9E2072689)

Received(April 25, 2022), Review Result(1st: May 10, 2022), Accepted(June 10, 2022), Published(June 30, 2022)



© 2022 The Authors. Published by NCIS.
This is an open access article licensed under the Creative Commons Attribution-NonCommercial 4.0 International License.
To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc/4.0/>.

causality. Also, with the development of artificial intelligence, it was successful to measure the distortion of the face, which is one of the symptoms from stroke patients. However, stroke can show other symptoms, such as voice tremors, in addition to facial paralysis. Therefore, this study proposes a deep learning model framework using artificial intelligence, focusing on the inarticulateness of speech and the distortion of the face among the symptoms of stroke. The proposed model applied transfer learning to improve accuracy and overcome limitations of insufficient datasets. As a result, the performance was improved by 0.7% in training accuracy, for validation, 13.9% in stroke accuracy, and 4.6% in general patients.

Keyword : Stroke, Mel-Spectrogram, Transfer Learning, Feature Extraction

1. 서론

4차 산업혁명이 도래하면서 다양한 산업과 연구 분야에 많은 발전이 이루어졌다. 헬스케어 분야의 경우, 인공지능, 빅데이터, 사물인터넷(IoT), 클라우드/엣지 컴퓨팅 등 정보통신기술(ICT)이 주목받고 있다 [1]. 예를 들어 IoT는 다양한 의료 기기와 의료 기관 시스템 간에 다양한 유형의 건강 및 의료 정보를 교환하게 해준다. 세계보건기구(WHO)에 따르면 세계 인구는 빠르게 고령화를 향하고 있다. 이들은 고령화로 인하여 의료 비용과 각종 만성 질환이 증가될 것으로 예상하고 있다.

세계보건기구에 따르면 2019년 세계에서 가장 큰 사망원인이 되는 허혈성 심장병은 전 세계 사망자의 16%를 차지한다. 뇌졸중은 그 다음 주요 원인으로 11%를 차지한다. 이는 2016년에 3위 인 것을 감안한다면 3년만에 2위를 차지한 것이다.

뇌졸중이란 뇌로 가는 혈류가 차단되어 뇌세포의 괴사를 일으키면서 발생하는 신경학적 증상을 말한다. 이는 뇌혈관 질환과 같은 표현으로 쓰이며, 한국에서는 ‘중풍’이란 말로 불려왔다. 현재 한국의 경우, 사망원인 1위가 뇌졸중인 정도로 많은 환자들이 앓고 있는 질환이다 [2-3].

뇌졸중은 다양한 연구와 임상 실험을 통해 여러 위험 요소가 보고되었다. 이는 발생 후에 치료를 하더라도 반신마비, 언어장애, 심하면 사망에 이를 수 있는 아주 위험한 병이다. 하지만 뇌졸중의 경우 질환 초기에 발견하게 된다면 그 중증 정도를 크게 하락시킬 수 있다 [4-6]. 또한 고혈압, 흡연, 당뇨병, 비만 등 조절 가능한 요소들을 환자 개개인이 적절히 관리하고 치료한다면 뇌졸중의 발생을 지연시키거나 차단시킬 수 있다 [7]. 미국의 경우, 국민건강영양조사(NHANES)에서 고혈압 환자를 대상으로 설문조사에 따르면, 약 30%의 사람들이 자신의 고혈압 여부를 인지하지 못하고 있었고, 고혈압 환자 중 15%의 환자들은 약물을 복용하지 않고 있다 [8]. 한국의 경우, 건강보험심사평가원에 의하면 2020년 673.1만명이 뇌졸중을 앓고 있었으며 이 수는 매해 증가하며 2016년 대비 약 12.3%가 증가하였다 [9].

이를 대비하여 각국은 질병 예방 시스템 체계를 구축하고 있다. 이에 따라 개인건강기록(PHR), 전자의무기록(EMR), 유전자 정보 등 개인의 의료데이터를 생성 및 수집되고 있다. 수년에 걸쳐 방대한 양의 의료 데이터가 저장되어 있지만, 그동안 제대로 활용되고 있지 못하였다.

최근 수년간 머신러닝 및 딥러닝 인공지능의 발전에 따라 다양한 분야의 데이터들을 적극적으

로 활용하고 있다. 의료분야에서도 이를 통해 과거 의사만 진단하는 방법과 달리, 빅데이터를 기반으로 학습한 인공지능이 보다 심층적인 분석을 하여 의사의 진단 결정에 도움을 주고 있다.

하지만 뇌졸중의 경우 한계점이 분명하게 존재한다. 기존에는 대부분의 환자들이 자신이 뇌졸중 환자인 것을 인지하지 못하고 있다는 점과, 뇌졸중 환자가 의사의 진단을 받기 위해서는 CT촬영과 같이 고비용 장시간의 투자가 필요했다. 이를 극복하기 위해 증상을 기반으로 간단한 초기진단 방법들을 만들어 CPSS 검사, FAST 검사, NIHSS 설문조사와 같이 간단한 검사를 통해 자가진단을 하거나 의사를 통한 진단이 진행됐다 [10]. 이때 인공지능의 발전으로 인해 의사가 눈으로 감지하지 못하는 미세한 영역까지 감지할 수 있게 되었다 [11].

또한 음성인식 등 음성 관련 연구의 경우 합성곱 신경망(Convolutional Neural Network, CNN), LSTM(Long Short-Term Memory)등의 모델과 이를 비롯한 연구 배경을 바탕으로 하나의 음성 파일에 있는 여러 특징값 중 한 가지를 분석하는 멜-스펙트로그램(Mel-Spectrogram) 혹은 MFCC(Mel Frequency Cepstral Coefficient)등의 방법들이 연구되었다.

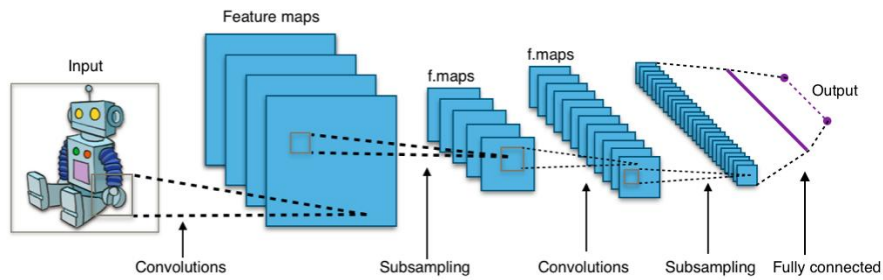
이에 본 연구에서는 뇌졸중 초기진단 방법 중 FAST검사의 특징에서 Face와 Specch에 초점을 맞춘 모델을 제안한다. 모델은 음성과 얼굴 이미지 각각의 특징값을 추출하여 뇌졸중 여부를 예측하는 딥러닝 모델 프레임워크이다. 제안된 모델은 224*224*3의 이미지를 입력으로 받아 EfficientNet으로 특징 벡터를 생성한다. 그다음 각각의 특징 벡터값을 하나의 벡터값으로 합친 후 마지막 신경망 모델로 최종 분류를 수행한다.

본 논문의 구성은 다음과 같다. 2장에서는 본 연구에 사용되는 배경지식(CNN, 전이학습, 멜-스펙트로그램)에 대하여 설명하고 뇌졸중 관련 기존 연구들을 기술한다. 3장에서는 데이터셋에 대한 설명과 전처리 방법 그리고 제안된 모델의 구조와 구현 방법에 대하여 설명한다. 4장에서는 실험에 대한 결과 분석과 연구 과정의 한계점 및 향후 연구 방향성에 대하여 논한다.

2. 관련연구

2.1. CNN

Convolutional Neural Networks(CNN)은 딥러닝에서 이미지 및 영상 데이터를 처리할 때 사용되는 모델이다 [12]. 이는 1차원 형태의 데이터만 취급 가능한 DNN과는 다르게, 이미지의 공간적 지역적 정보를 유지한 채 특성들의 계층을 생성해준다. 예를 들어 [그림 1]에서는 CNN 모델을 통해 로봇의 팔에 대한 특성들을 나타내준다. 이 모델의 주 기능은 이미지의 일부분을 분석하여 각 픽셀 간의 연관성을 수치화하여 나타내는 것이다.

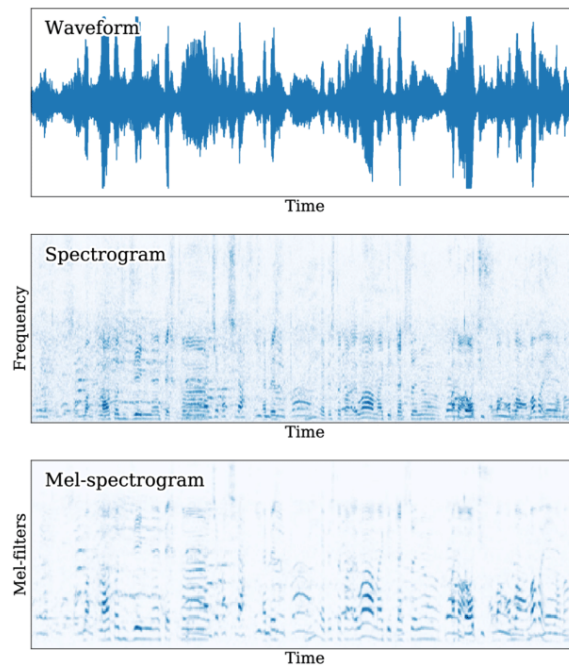


[그림 1] CNN 구조

[Fig. 1] CNN Architecture

2.2. 멜-스펙트로그램

멜 스펙트로그램이란 Mel 스케일로 변환된 스펙토그램이다 [13]. 스펙토그램이란, 신호의 주파수 영역을 시각화한 것으로 주파수의 범위를 [그림 2]와 같이 이미지 형태로 나타낸다.



[그림 2] Mel-Spectrogram으로 음성을 변환한 결과

[Fig. 2] Result of Voice changed by Mel-Spectrogram

Mel-Scale이란 주파수의 일부를 비선형으로 변환한 결과이다. 다시 말해 소리의 크기 단위(Hz)를 빈도수로 변환하여 특정 영역의 주파수가 반복되는 부분을 표현해준다. 보통 음성 신호를 분석할

때 신호는 주파수(Frequency)와 강도(Amplitude)로 나뉘어 각 신호의 세기를 알 수 있다. 그러나 이 과정에서 시간에 관련된 정보가 사라지는 한계점이 있다. 이는 신호를 일정한 크기의 창(window)로 나누어 하나씩 변환을 적용하면 해결할 수 있다. 결과적으로 각 신호에 대한 크기를 시간, 주파수, 강도로 분해하여 정보를 획득할 수 있다. 이 방법은 자동 소리 인식, 소리 식별, Text-To-Speech(TTS), Speech-To-Text(STT) 등 다양한 분야에 활용되고 있다 [14-16].

2.3. 전이학습

전이학습(Transfer Learning)은 특정 분야에서 학습된 신경망의 일부분을 유사하거나 새로운 분야의 신경망 학습에 활용되는 것을 말한다. 예를들어 ResNet [17], VGG [18]와 같은 신경망은 CNN의 특징인 이미지 추출 기능을 활용하여 사전에 규모가 매우 큰 데이터셋으로 학습을 한다. 이를 활용하는 데이터셋에 맞추어 파인튜닝을 하여 개선된 성능을 확인할 수 있다. 따라서 전이학습은 학습 데이터의 수가 적을 때도 효과적이며, 학습 속도와 정확도가 다른 모델에 비하여 높다.

2.4 인공지능 기술을 활용한 뇌졸중 감식

기존의 뇌졸중 인식 연구들은 얼굴에 나오는 증상 중 비대칭성과 입의 처진 정도의 감도를 측정하여 판정하거나 [19][20], 웨어러블 장치를 활용한 연구들이 있다 [21]. 안면인식을 통한 뇌졸중 감지의 경우, 얼굴의 랜드마크 추적을 통해 비대칭성을 파악한다 [22][23]. 그러나 안면만으로는 주관적, 정량적 평가를 하기에 특징점을 잡아내지 못하는 한계점들이 있다. 따라서 인간의 눈으로는 할 수 없는 미세하고 정확한 위치 파악을 위해 머신러닝 알고리즘을 개발하기도 하였다 [24].

딥러닝의 급진적인 발달과 지속해서 쌓이는 데이터로 딥러닝을 활용한 연구들이 증가하고 있다. SRP모델을 설계하기 위한 연구 [25], 머신러닝·딥러닝 기반으로 뇌졸중 예측을 하는 연구들이 있다 [26][27]. 또한 전이학습의 발전으로 인해 데이터가 적은 경우에도 모델을 학습할 수 있고, 그 성능들이 우수한 결과를 가진 연구들이 진행되어왔다.

이에 본 연구에서는 멜-스펙트로그램으로 음성을 이미지로 변환시켜, 음성과 얼굴 이미지의 각각의 특성들을 전이학습을 활용해 벡터값으로 추출한다. 그 다음 하나의 벡터값으로 합친 후 뇌졸중을 판별하는 모델을 제안한다.

3. 모델 프레임워크

3.1 데이터셋 및 전처리

본 연구에 사용된 데이터는 2017.09월부터 4년간 길병원에 내원한 환자 데이터로, 총 319명의 환

자의 얼굴 데이터와 음성 데이터로 이루어져 있다. 전체 환자 중 290명은 뇌졸중 환자, 29명은 일반 내원 환자이다. 얼굴 데이터의 구성은 [“Cheek”, “Close_eyes”, “Front”, “Left”, “Raise_eyebrows”, “Right”, “Teeth”] 으로 각각 [“볼을 부풀림”, “눈을 감음”, “정면 응시”, “왼쪽 응시”, “눈썹 들기”, “우측 응시”, “이빨 드러냄”]이다. 음성 데이터의 경우 FAST 뇌졸중 검사 중 Speech 에 해당하는 것으로, 특정 문장을 시간제한 없이 읽도록 하여 녹음한 데이터이다. 모든 환자들은 간호사의 지도에 따라 녹음이 된다. 그 과정에서 간호사의 목소리, 주변 잡음 등이 함께 녹음되기 때문에 섞인 잡음을 noisereduce 라이브러리를 사용하여 제거해주어야 한다. 이는 음성 파일에서 각 진폭을 측정하여 에너지 값을 계산하고(1), 그에 대한 평균 제곱근 값을 구하여(2) 불필요한 신호의 크기를 줄이는 방법이다. 이때 n 은 각 신호를, N 은 음성파일의 전체 길이를 의미한다.

$$ENERGY = \sum_{n=1} |x(n)|^2 \quad (1)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_n |x(n)|^2} \quad (2)$$

3.2. 멀티 모델 프레임워크

본 연구에서는 [그림 3] 과 같이 일반적인 Conv블록을 사용하여 학습을 하는 경우 과적합이 나와 사용할 수 없었다. 과적합이 일어나는 원인으로 데이터셋의 불균형과 부족한 양으로 판단했다. 따라서 데이터셋이 부족한 경우에도 활용할 수 있는 사전학습 된 전이학습 모델을 활용하였다. 기존의 연구들을 보면 ConvNet의 성능을 향상시키기 Scaling up 시도를 하였다.

대표적으로 ResNet, ResNet-200과 같이 모델의 레이어 개수를 늘리거나, channel width를 늘려 정보를 더 많이 담거나, 입력 이미지의 resolution, 즉 input image의 해상도를 높여 성능 향상을 이뤄냈다. 이후 다양한 모델의 발전으로 VGGNet, GoogleNet과 같은 모델들을 활용하여 더 좋은 성능을 만든다는 것을 확인할 수 있었다. 하지만 본 연구의 경우 수많은 블록과 파라미터로 인해 오히려 과적합을 이뤄내는 것을 알 수 있었다. 이에 적은 레이어층과 파라미터 개수를 가지고도 높은 성능을 낼 수 있는 EfficientNet [28]을 활용하였다. EfficientNet은 mobile inverted bottleneck convolution (MBConv)을 사용한다. 이는 MobileNetV1 [29]에서 제안한 Depthwise separable Conv와 squeeze and excitation network(SENet)에서 제안된 각 채널들의 중요 정보만 추출해 가져가는 방법 2개를 활용한 것이다. 따라서 제안된 모델은 EfficientNet을 활용하여 얼굴 이미지와 음성 각각의 특성 추출을 한 모델로 뇌졸중을 판별하는 모델을 제안하였다. 이 모델을 통해 학습된 데이터는 마지막 완전 연결 계층을 통해 특징 벡터값이 추출된다. 본 연구에서는 EfficientNet-B0로 진행했는데, 다른 B1~B7의 경우 상대적으로 모델의 연산량이 많아 쉽게 과적합되기 때문이다. [그림 4]은 제안된 모델의 구조

4. 실험 및 평가

본 연구에 사용된 라이브러리 및 개발환경은 [표 1]과 같다.

[표 1] 개발환경

[Table 1] Environment

개발환경	
Python	3.8
Numpy	1.19.5
Tensorflow	2.8
Keras	2.8.0
Imutils	0.5.4
Librosa	0.9.1
GPU	Nvidia 1080

인구의 고령화로 인하여 인간의 가장 큰 사망원인으로 뇌졸중이 급부상하고 있다. 초기에 치료를 한다면 중증 정도를 크게 감소가능한 질환이기 때문에, 이런 심각성을 인지하고 치료 및 검출을 하기위한 다양한 연구가 진행되고 있다. 그러나 뇌졸중을 갖고 있는 환자들은 본인이 인지하지 못하는 경우가 대다수이며, 약물을 적시에 복용하고 있지 않다는 설문조사 결과가 있었다. 의료진은 고비용 장시간의 투자가 필요한 CT촬영 대신 나타나는 증상을 기반으로 초기 진단 방법들을 만들었다. 또한 하드웨어의 발전과 데이터의 방대한 양으로 머신러닝 및 딥러닝을 활용한 연구가 진행되어왔다. 기존 연구에는 얼굴의 비대칭 여부 혹은 일그러짐 정도를 측정하여 뇌졸중 여부를 판단하였다. 하지만 의료 데이터의 경우 아직 공개적으로 공개된 데이터가 없을 뿐만 아니라, 해당 데이터를 보관하는데도 문제가 많다. 그러나 기술과 모델의 발전으로 전이학습을 활용하여 데이터셋이 부족한 연구에도 다양한 실험이 이루어져왔다. 본 연구에서는 뇌졸중의 안면의 일그러진 정도에 초점이 맞추어져 있던 기존 연구에서, 음성을 함께 활용하는 멀티 모델 프레임워크를 제안했다. 이를 비교평가하기 위해 동일한 EfficientNet 모델로 각각의 데이터셋을 적용한 모델을 만들었다. 그에 대한 결과는 [표 2]에 나타냈다

실험결과를 보면 본 연구에서 제안된 멀티 모델 프레임워크가 일반 음성 데이터만을 사용했을 때 보다 뇌졸중 환자 검증 과정에서 성능이 14%, 일반 환자 검증 과정에서 4.6%가 개선된 것을 알 수 있다. 현재 보유하고 있는 음성 데이터의 경우, 휴대폰으로 병원에서 녹음이 되었기 때문에 잡음이 많이 포함되어 있다. 따라서 전처리 과정에서 오히려 학습에 필요한 특성들이 삭제되면서 상대적으로 낮은 검증 정확도를 보이는 것으로 예상된다. 본 연구에서 제안된 모델의 경우, 얼굴

이미지의 데이터를 특성으로 추가하여 학습을 진행했기 때문에 개선된 성능을 보였다.

[표 2] 모델 훈련 결과

[Table 2] Model Results

EfficientNet 활용한 모델			
모델 종류		Model 1. 음성	Model 2. 얼굴 이미지+음성
훈련 정확도(%)		73.7	74.4
F1-Score	Stroke	58. 1	72
	Normal	69.2	73.8

5. 결론

본 연구는 복합생체 정보를 인공지능으로 학습하여 뇌졸중을 감식하는 모델을 제안하였다. 음성 데이터만으로 학습한 모델보다 복합생체 정보를 이용하여 학습한 모델이 성능이 우수한 것을 확인할 수 있었다. 다양한 특징을 가진 데이터를 통해 더 정확한 예측을 할 수 있다는 것을 검증하였다. 그러나 인공지능 연구에서 가장 중요하게 고려되어야 할 것은 데이터이다. 균등한 비율의 데이터와 수많은 양의 데이터가 필요하다. 연구를 시작하며 실제 뇌졸중 환자를 대상으로 데이터를 측정하고자 하였으나, 병원에 내원하는 부서가 다르므로 시진환자와 청진환자의 데이터를 동시에 측정하기에 어려움이 있었다. 또한, 음성 데이터의 경우, 뇌졸중 환자임에도 불구하고 젊은 층에 속하여 상대적으로 말을 뚜렷하게 하지만 일반 환자임에도 불구하고 말을 어눌하게 하는 환자들도 있었다. 그뿐만 아니라 방음이 되는 환경에서 녹음을 진행하여 부가적인 잡음이 포함되지 않도록 할 필요성이 있었다. 향후 과제에서는 데이터의 수집 과정과 환자들의 특성을 미리 파악하고, 정제된 데이터를 확보하고 연구를 진행한다면 더 높은 성능을 기대할 수 있을 것으로 생각한다.

References

- [1] O. O. Ajayi, A. B. Bagula, K. MA, "Fourth Industrial Revolution for Development: The Relevance of Cloud Federation in Healthcare Support", IEEE Access, vol. 7, December 2019, pp. 185322-185337, doi: 10.1109/ACCESS.2019.2960615.
- [2] A. Subudhi, M. Dash, S. Sabut, "Automated segmentation and classification of brain stroke using expectation-maximization and random forest classifier", Biocybernetics and Biomedical Engineering, vol. 40, no. 1, May 2019, pp. 277-289, doi: 10.1016/j.bbe.2019.04.004.
- [3] H. J. Lee, J. S. Lee, J. C. Choi, Y. J. Cho, B. J. Kim, H. J. Bae, D. E. Kim, W. S. Ryu, J. K. Cha, D. H. Kim, H. W. Nah, K. H. Choi, J. T. Kim, M. S. Park, J. H. Hong, S. L. Sohn, K. Kang, J. M. Park, W. J. Kim, J. Lee, D. L. Shin, M. J. Yeo, K. B. Lee, J. G. Kim, S. J. Lee, B. C. Lee, M. S. Oh, K. H. Yu, T. H. Park, J. Lee, K. S. Hong, "Simple Estimates of Symptomatic Intracranial Hemorrhage Risk and

- Outcome after Intravenous Thrombolysis Using Age and Stroke Severity”, *Journal of Stroke*, vol. 19, no. 2, May 2017, pp. 229-231, doi: 10.5853/jos.2016.01109.
- [4] Y. D. Kim, Y. H. Jung, G. Saposnik, “Traditional Risk Factors for Stroke in East Asia”, *Journal of Stroke*, vol. 18, no. 3, September 2016, pp. 273-285, doi: 10.5853/jos.2016.00885.
- [5] M. H. F. Poorthuis, A. M. Algra, A. Algra, L. J. Kappelle, C. J. M. Klijn, “Female- and Male-Specific Risk Factors for Stroke: A Systematic Review and Meta-analysis”, *JAMA Neurology*, vol. 74, no. 1, January 2017, pp. 75-81. doi: 10.1001/jamaneurol.2016.3482.
- [6] Y. V. Kalkonde, S. Alladi, S. Kaul, V. Hachinski, “Stroke Prevention Strategies in the Developing World”, *Stroke*, vol. 49, no. 12, December 2018, pp. 3092-3097, doi: 10.1161/STROKEAHA.118.017384.
- [7] A. K. Boehme, C. Esenwa, M. S. Elkind, “Stroke Risk Factors, Genetics, and Prevention”, *Circulation Research*, vol. 120, no. 3, February 2017, pp. 472-495, doi: 10.1161/CIRCRESAHA.116.308398.
- [8] Centers for Disease Control and Prevention, “1999-2000 Data Documentation, Codebook, and Frequencies Blood Pressure & Cholesterol (BPQ)”, *cdc.gov*, <https://wwwn.cdc.gov/Nchs/Nhanes/1999-2000/BPQ.htm>, (accessed September 27, 2020).
- [9] Salary Information Management Department, “Statistics on diseases and medical treatment in daily life”, <https://hira.or.kr>, <https://www.hira.or.kr/bbsDummy.do?pgmid=HIRAA020045010000&brdScnBltno=4&brdBltno=2361&pageIndex=1>, (accessed March 16, 2022).
- [10] J. Harbison, O. Hossain, D. Jenkinson, J. Davis, S. J. Louw, G. A. Ford, “Diagnostic Accuracy of Stroke Referrals From Primary Care, Emergency Room Physicians, and Ambulance Staff Using the Face Arm Speech Test”, *Journal of Stroke*, vol. 34, no. 1, December 2002, pp. 71-76, doi: 10.1161/01.STR.0000044170.46643.5E.
- [11] M. S. Sirsat, E. Ferme, J. Camara, “Machine Learning for Brain Stroke: A Review”, *Journal of Stroke and Cerebrovasc Diseases*, vol. 29, no. 10, October 2020, doi: 10.1016/j.jstrokecerebrovasdis.2020.105162.
- [12] A. Krizhevsky, I. Sutskever, G. E. Hinton, “ImageNet classification with deep convolutional neural networks”, *Journal of Communications of the ACM*, vol. 60, no. 6, June 2017, pp. 84-90. doi: 10.1145/3065386.
- [13] B. Zhang, J. Leitner, S. Thornton, “Audio Recognition using Mel Spectrograms and Convolution Neural Networks”, *Computer Science*, 2019.
- [14] J. Shen, R. Pang, R. J. Weiss, M. Schuster, N. Jaitly, Z. Yang, Z. Chen, Y. Zhang, Y. Wang, R. Skerry-Ryan, R. A. Saurous, Y. Agiomyriannakis, Y. Wu, “Natural TTS Synthesis by Conditioning Wavenet on MEL Spectrogram Predictions”, *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, April 15-20, 2018, Calgary, Canada, pp. 4779-4783, doi: 10.1109/ICASSP.2018.8461368.
- [15] T. Tran, J. Lundgren, “Drill Fault Diagnosis Based on the Scalogram and Mel Spectrogram of Sound Signals Using Artificial Intelligence”, *Journal of IEEE Access*, vol. 8, November 2020, pp. 203655-203666, doi: 10.1109/ACCESS.2020.3036769.
- [16] D. de Benito-Gorron, A. Lozano-Diez, D. T. Toledano, J. Gonzalez-Rodriguez, “Exploring convolutional, recurrent, and hybrid deep neural networks for speech and music detection in a large audio dataset”, *Journal of Audio Speech Music Processing*, vol. 9, June 2019, doi: 10.1186/s13636-019-0152-1.
- [17] K. He, X. Zhang, S. Ren, J. Sun, “Deep Residual Learning for Image Recognition”, 2016 IEEE Conference

- on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, USA, pp. 770-778, doi : 10.1109/CVPR.2016.90.
- [18] K. Simonyan, A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition", The International Conference on Learning Representations (ICLR), May 7-9, 2015, San Diego, USA, doi: 10.48550/arXiv.1409.1556.
- [19] S. Umirzakova, A. Abdusalomov, T. K. Whangbo, "Fully Automatic Stroke Symptom Detection Method Based on Facial Features and Moving Hand Differences", International Symposium on Multimedia and Communication Technology (ISMAC), August 19-21, 2019, Quezon City, Philippines, pp. 1-5, doi: 10.1109/ISMAC.2019.8836166.
- [20] C. -Y. Chang, M. -J. Cheng, M. H. -M. Ma, "Application of Machine Learning for Facial Stroke Detection", IEEE 23rd International Conference on Digital Signal Processing (DSP), November 19-21, 2018, Shanghai, China, pp. 1-5, doi: 10.1109/ICDSP.2018.8631568.
- [21] M. Kos, J. Zenko, D. Vlaj, I. Kramberger, "Tennis stroke detection and classification using miniature wearable IMU device", International Conference on Systems, Signals and Image Processing (IWSSIP), June 1-3, 2016, Sofia, Bulgaria, pp. 1-4, doi: 10.1109/IWSSIP.2016.7502764.
- [22] G. S. Parra-Dominguez, R. E. Sanchez-Yanez, C. H. Garcia-Capulin, "Facial Paralysis Detection on Images Using Key Point Analysis", Applied Sciences, vol. 11, no. 5, March 2021, doi: 10.3390/app11052435.
- [23] O. Foong, K. Hong, S. Yong, "Droopy Mouth Detection Model in stroke warning", 3rd International Conference on Computer and Information Sciences (ICCOINS), August 15-17, 2016, Kuala Lumpur, Malaysia, pp. 616-621, doi: 10.1109/ICCOINS.2016.7783286.
- [24] D. L. Guarin, Y. Yunusova, B. Taati, J. R. Dusseldorp, S. Mohan, J. Tavares, M. M. Veen, E. Fortier, T. A. Hadlock, N. Jowett, "Toward an Automatic System for Computer-Aided Assessment in Facial Palsy", Journal of Facial Plastic Surgery & Aesthetic Medicine, vol. 22, no. 1, February 2020, pp. 42-49, doi: 10.1089/fpsam.2019.29000.gua.
- [25] A. Khosla, Y. Cao, C. C. Lin, H. Chiu, J. Hu, H. Lee, "An integrated machine learning approach to stroke prediction", In Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining (KDD '10). Association for Computing Machinery, July 25-28, 2010, New York, USA, pp. 183-192. doi: 10.1145/1835804.1835830.
- [26] J. Chen, Y. Chen, J. Li, J. Wang, Z. Lin, A. K. Nandi, "Stroke Risk Prediction With Hybrid Deep Transfer Learning Framework", IEEE Journal of Biomedical and Health Informatics, vol. 26, no. 1, January 2022, pp. 411-422, doi: 10.1109/JBHI.2021.3088750.
- [27] J. Yu, S. Park, S. Kwon, C. M. B. Ho, C. Pyo, H. Lee, "AI-Based Stroke Disease Prediction System Using Real-Time Electromyography Signals", Journal of Applied Sciences, vol. 10, no. 19, September 2020, pp. 6791, doi: 10.3390/app10196791.
- [28] M. Tan, Q. V. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks", 36th International Conference on Machine Learning, June 9-15, 2019, California, USA, pp. 6105-6114, doi: 10.48550/arXiv.1905.11946.
- [29] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, H. Adam, "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications", arXiv preprint arXiv 2017, April 2017, doi: 10.48550/arXiv.1704.04861