creative commons
COMMONS DEED

저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:

저작자표시. 귀하는 원저작자를 표시하여야 합니다.

비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.

변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 이용허락규약(Legal Code)을 이해하기 쉽게 요약한 것입니다.

Disclaimer

dCollection

치의과학석사 학위논문

딥러닝 분석을 활용한
치아우식 탐지의 신뢰도 평가

2019년 8월

서울대학교 대학원

치의과학과 예방치과학전공

김 선 집

딥러닝 분석을 활용한
치아우식 탐지의 신뢰도 평가

지도교수 조 현 재

이 논문을 치의과학석사 학위논문으로 제출함

2019년 5월

서울대학교 대학원

치의과학과 예방치과학전공

김 선 집

김선집의 치의과학석사 학위논문을 인준함

2019년 7월

위 원 장 진 보 형 (인)

부위원장 조 현 재 (인)

위 원 한 동 헌 (인)

국문초록

딥러닝 분석을 활용한 치아우식 탐지의 신뢰도 평가

서울대학교 대학원 치의과학과 예방치과학전공

(지도교수 : 조 현 재)

치아우식의 검진 도구는 탐침 및 치과 방사선 사진 영상이 주로 사용되고 있으나, 최종적으로는 진단자의 경험에 의존하는 경향이 있으며, 진단자 간 신뢰도가 만족스럽지 않은 경우도 있다. 이러한 문제를 보완하고자 최근 영상 분석에 우수한 성능을 보이는 딥러닝을 활용하여 영상을 처리하거나 패턴을 인식하는 진단 기술을 치과 분야에 적용하여 치아우식의 진단 정확성을 높이고자 하였다.

이번 연구는 인터넷 검색 엔진에서 치아 영상를 수집하는 Search engine scraping 방법을 사용하여, 서울대학교 치의학대학원 연구윤리심의위원회의 승인(S-D20190007)을 받아 2019년 4월부터 2019년 5월까지 약 2개월에 걸쳐 진행되었다. 치아 영상은 세계보건기구 지침에 의해 건전치아, 우식치아, 처치치아로 나누었다. 최종적으로 각각 70개의 영상을 수집하였으며, 딥러닝

모델을 학습시키기 위한 훈련 세트 50개와 이를 검증하는 시험세트 20개를 무작위로 배치하였다. 건전치아와 우식치아를 분류하는 모델에서 신뢰도를 확인하기 위해 ROC curve, AUC, 일치도, 민감도, 특이도, 양성예측도, 음성예측도를 산출하였고, 건전치아, 우식치아, 처치치아를 분류하는 모델에서는 각 분류에 따른 일치도를 확인한 결과, 다음과 같은 결론을 얻었다.

1. 치과 영상의 특징을 추출하여 신뢰도 높은 모델을 구축하기 위해서는 깊은 신경망으로 설계되어야 한다.
2. 건전치아와 우식치아를 분류하는 모델에서, 전이학습 적용모델은 전이학습 비적용 모델(50.0%)에 비하여 월등히 높은 일치도를 나타내었다(90.0%, AUC:0.963).
3. 건전치아, 우식치아, 처치치아를 분류하는 모델에서, 일치도는 감소하였고(61.7%), 각 영상의 일치도는 처치치아(80.0%), 우식치아(35.0%), 건전치아(25.0%) 순이었다.

주요어 : 딥러닝, 인공지능, 치아우식, 합성곱 신경망

학번 : 2017-29616

목 차

제 1장 서론 ---------- 1

제 1절 연구 배경 및 목적 ---------- 1

제 2장 인공지능의 이해 ---------- 3

제 1절 인공지능의 역사 ---------- 3

제 2절 인공지능의 개념 ---------- 4

제 3절 Convolutional neural network ---------- 9

제 4절 딥러닝 기술 용어 ---------- 14

제 5절 프로그래밍 언어 ---------- 20

제 3장 연구대상 및 방법 ---------- 22

제 1절 연구 설계 및 대상 ---------- 22

제 2절 연구 방법 ---------- 24

제 4장 연구 결과 ---------- 26

제 1절 두가지 범주를 분류하는 모델 ---------- 26

제 2절 세가지 범주를 분류하는 모델 ---------- 32

제 3절 전이학습 적용 모델과 비적용 모델 비교 ---------- 35

제 5장 고안 ---------- 36

제 6장 결론 ---------- 40

제 1절 연구 결론 ---------- 40

제 2절 향후 연구 ---------- 41

참 고 문 헌 ---------- 42

Abstract ---------- 45

표 목 차

[Table 2-1] Python library for deep learning analysis ----------21

[Table 3-1] Inclusion/exclusion criteria for dental image ----------23

[Table 4-1] Confusion matrix based on results of dental caries classification using CNN model ----------30

[Table 4-2] Accuracy, precision, sensitivity, specificity, positive predictive value, and negative predictive value for the detection of dental caries ----------30

[Table 4-3] Accuracy of CNN model according to tooth condition of clinical image ----------34

그 림 목 차

[Fig. 2-1] Comparison between a human neuron and an ANN synapses -6

[Fig. 2-2] An example of neural network ----------7

[Fig. 2-3] Difference in image recognition between human and computer-9

[Fig. 2-4] Back propagation ----------10

[Fig. 2-5] Difference between kernel size and filter size ----------11

[Fig. 2-6] An example of max pooling ----------12

[Fig. 2-7] The formula of cross entropy ----------14

[Fig. 2-8] Gradient descent ----------15

[Fig. 2-9] Global minimum and local minimum ----------16

[Fig. 2-10] Sigmoid function ----------17

[Fig. 2-11] ReLU function ----------18

[Fig. 2-12] Softmax function ----------19

[Fig. 3-1] Research flow chart ----------22

[Fig. 3-2] Preprocessing: pixel size resizing ----------24

[Fig. 4-1] Transfer learning non-use model ----------26

[Fig. 4-2] VGG16 model: non-trainable parameter ----------28

[Fig. 4-3] Trainable parameter in transfer learning ----------29

[Fig. 4-4] Loss and accuracy by epoch in two classification model -----29

[Fig. 4-5] Receiver operating characteristic curve ----------31

[Fig. 4-6] Loss and accuracy by epoch in three classification model ----33

[Fig. 4-7] Comparison of accuracy according to application of transfer learning -35

[Fig. 4-8] Comparison of loss value according to application of transfer learning -35

용어의 정의

1.	인공지능 (artificial intelligence, AI)	인공지능은 컴퓨터(기계)가 인간과 같은 지능을 갖도록 만드는것을 의미
2.	머신러닝 (machine learning)	인공지능의 한 분야로, 패턴과 추론에 의존하지 않고 효과적으로 특정 작업을 수행하기 위해 컴퓨터 시스템이 사용하는 알고리즘 및 통계 모델에 대한 과학적인 연구를 의미
3.	딥러닝(deep learning)	머신러닝의 한 분야로, 비선형 변환기법의 조합을 통해 높은 수준의 추상화를 시도하는 알고리즘의 집합을 의미
4.	인공신경망(artificial neural network, ANN)	머신러닝과 인지과학에서 생물학의 신경망에서 영감을 얻은 통계학적 학습 알고리즘
5.	Convolutional neural network (CNN)	주로 시각 영상을 분석하는 데 사용되는 인공신경망의 한 종류
6.	전이학습 (transfer learning)	머신러닝에서 하나의 문제를 해결하면서 얻은 지식을 저장하여, 다른 문제에 적용하는 방식을 의미
7.	Parameter	함수의 정의에서 전달받은 인수를 함수 내부로 전달하기 위해 사용하는 변수를 의미

제 1장 서 론

제 1 절 연구 배경 및 목적

치아우식의 치료법은 지난 수십년 동안 지속적으로 발전해왔으며, 다양한 수복 방법이 성공적으로 개발되었으나 치아우식을 진단하는 방법론은 아직 크게 개선되지 않았다[1]. 특히, 치아의 깊은 소와 열구, 긴밀한 치아 인접면 및 이차 병변이 존재할 때, 초기 우식을 발견하기가 어렵다. 결국은 많은 병변에서 치아우식이 상당이 진행이 되었을 때 발견되는 경우가 있다. 현재 정확한 치아우식의 진단을 위해 신뢰할 수 있는 진단 도구로는 치과 방사선 촬영, 탐침 등을 있으며 최종적으로는 진단자의 경험적 증거에 의존하는 경향이 있다[2]. 이러한 이유로 치과의사의 개인적인 견해와 진단 방법으로 동일한 대상자라도 우식치아의 개수가 달라질 수 있으며, 치료의 정도가 달라질 수 있다는 문제가 야기될 수 있다[3].

제4차산업 혁명으로 인공지능(artificial intelligence)에 대한 관심도가 높아지고 있으며, 기존의 정보통신기술과 융합되어 여러 분야에서 활용하기 위해 연구가 지속되고 있다[4]. 이러한 인공지능을 활용한 혁신적인 변화는 사물 인터넷(internet of things), 빅데이터(big data), 클라우드 컴퓨터(cloud computer), 3D printing 기술 등의 발전을 이루었으며 계속해서 활용 분야를 넓혀가고 있다. 인공지능의 발전은 업무 효율성을 향상 시키며, 인간이 범할 수 있는 오류를 줄여줌으로서 인간이 작업해왔던 업무의 형태가 수작업이 아닌 디지털 업무(digital labor)들로 대체되고 있다[5]. 이 흐름은 최근에 급진적으로 발생해오던 현상이 아닌, 과학의 발전으로부터 직업의 역사성과 창조성의 특성을 띄며 변화해 왔다.

의료분야도 예외는 아니며, 지난 20년 동안 컴퓨터 의료진단(computer-aided diagnosis)은 디지털 의료 데이터의 접근성 증가, 연산 능력의 향상 및 인공지능의 발전으로 인해 크게 발전해 왔다[6]. 이러한 발전은 의료 전문가를 대체 하는 것이 아니라 오히려 진단의 정확성을 높여 주는 하나의 도구로서 채택되었다. 특정 의료분야뿐만 아니라 점차 다양한 의료분야까지 인공지능을 활용하여 진단의 신뢰도를 높이고자 하는 움직임을 보이고 있으며, 치과 분야에서도 치근단 방사선 촬영 영상에서 치아우식을 진단하는 딥러닝 모델이 우수한 성능을 보인다고 보고되었다[2]. 최근 방사선 촬영 영상에서 딥러닝 모델을 적용하여, 치아우식의 진단 이외에도 치근단 병소, 상악동염, 수직파절 치아 진단 등 활발하게 연구가 진행되고 있다[7-9]. 방사선 촬영 영상은 환자가 치과에 내원하여 촬영해야 한다는 단점이 있으며, 대규모 구강검진을 실시 할때는 적합하지 않다. 그러나, 일반 카메라를 활용하여 구강내 치아 사진을 촬영한다면 비교적 쉽게 치아 영상을 수집할 수 있을 뿐만 아니라 접근성과 활용성 또한 우수할 것으로 판단하였다. 아직까지 치아 사진을 활용한 딥러닝 모델의 활용 가능성을 연구한 보고가 거의 없는 실정이며, 방사선 촬영 사진에서 딥러닝 분석의 한계점을 보완하고자 이번 연구에서 치과 임상 사진을 활용하여 딥러닝 분석을 적용한 치아우식 탐지의 신뢰도를 평가하고자 하였다.

제 2 장 인공지능의 이해

제 1절 인공지능의 역사

1. 인공지능의 탄생

인공지능의 역사는 고대시대부터 이루어졌으며 고전 철학자에 의해 인간의 사고 과정을 기계적으로 묘사하려고 시도해왔으나[10,11], 본격적으로 1940년대와 50년대에 다양한 분야의 과학자가 인공두뇌의 가능성을 논의하기 시작하였다[12].

2. 황금기

인공지능은 1956년 여름 다트머스 대학에서 열린 워크샵에서 하나의 연구분야로 설립되었다[13]. 당시에 기계에 의한 지적행동이 전혀 가능하지 않다고 믿었던 사람들의 통념을 깨고 컴퓨터가 대수학의 문제를 풀고 기하학의 정리를 증명하여 언어의 구조를 이해하기 시작하였다. 연구원들은 20년 이내에 완벽한 지능을 갖는 기계가 구축될 것이라고 낙관적으로 전망하면서 인공지능 개발에 투자하였다.

3. 암흑기

낙관적인 전망은 오래가지 못하였고, 여러 가지 난관에 봉착하게 된다. 영국의 저명한 수학자인 제임스 라이트힐(James Lighthill)이 강력하게 인공지능에 대한 비판[14]과 의회에 지속적인 압력에 의해 미국과 영국 정부는 더 이상 인공지능 연구를 지원하지 않게 되었다. 또한 컴퓨터 성능이 충분한 메모리와 처리속도를 구현해 낼 수 없었기 때문에 제한점이 많았다.

4. 빅데이터의 시대

21세기 초반에 빅데이터의 구축과 고성능 컴퓨터가 개발됨에 있어 인공지능 기술이 경제 분야에 성공적으로 적용되었다. 그 결과 2016년 인공지능 관련 시장이 80억 달러를 돌파하였다[15]. 인공지능 기술중 하나인 convolutional neural network (CNN)과 recurrent neural network (RNN)이 발전하게 되면서 영상, 문자, 음성 인식 등의 연구가 활발히 진행되었다[4].

제 2 절 인공지능의 개념

1. 인공지능(Artificial Intelligence, AI)

인공지능은 컴퓨터(기계)가 인간과 같은 지능을 갖도록 만드는것을 의미한다[12]. 인공지능을 설명할 때 2개의 하위 개념으로 설명할 수 있는데 '강한 인공지능'과 '약한 인공지능'이다.

'강한 인공지능'은 일반적으로 정의 내려진 개념과 비슷하게 인공적인 기계와 소프트웨어가 인간의 지능과 같은 방식으로 운영되는 시스템을 말한다[16]. 예를들어 인간의 질문 의도를 완벽하게 파악하여 대화를 나누거나, 한가지 분야에 국한되지 않고 여러 분야를 접목하여 사고 할 수 있는 지능을 의미한다. '강한' 이라는 의미는 '범용적'이라는 의미와 상통하며, 일반적인 인공지능과 구분을 지어 '인공 범용 지능(artificial general intelligence)' 이라고도 불린다.

반면 '약한 인공지능'은 인간이 할 수 있는 지적 활동을 효율적으로 수행하기 위해 지능이 작용하는 일부 의학적, 논리적 기전을 인간이 활용하는 시스템을 말한다[16]. '약한 인공지능'의 특징은 기계의 작동 구현 방식이 인간의 지능과는 근본적으로 다르다는 것을 인정한다. '약한 인공지능'은 '강한 인공지

능'을 실현하기 위한 방법중 하나이며, 일상생활에서 볼 수 있는 대부분의 인공지능은 '약한 인공지능'이라고 볼 수 있다. 예를 들어 문자 인식, 음성 인식, 통계 소프트웨어 등을 들 수 있다.

2. 머신러닝(machine learning)

머신러닝은 인공지능의 하위 집합으로 간주되며, 패턴과 추론에 의존하지 않고 효과적으로 특정 작업을 수행하기 위해 컴퓨터 시스템이 사용하는 알고리즘 및 통계 모델에 대한 과학적인 연구이다[17]. 머신러닝은 예측하고자 하는 결과값을 얻기 위해 단순하며, 명확한 알고리즘으로 구성되어 있는 것이 아닌 훈련세트(training set) 라고 하는 샘플 데이터(sample data)의 학습을 기반으로 한 복잡한 수학 모델이다. 훈련세트로 학습이 진행된 머신러닝 알고리즘 모델은 시험세트(test set)를 통해서 일치도, 학습률, 결과의 오차(loss) 등을 확인할 수 있다.

3. 인공신경망(Artificial Neural Network, ANN)

인공신경망은 머신러닝과 딥러닝의 핵심적인 작동방법으로 뉴런 작동기전에서 영감을 얻은 통계학적 학습 알고리즘이다. 뉴런의 구조는 신호를 받아들이는 가지돌기와 이를 전달하는 축삭, 그리고 다른 뉴런으로 신호를 전달할 수 있는 축삭말단으로 구성되어 있다. 인접한 두 뉴런이 신호를 전달할 수 있도록 시냅스라는 구조가 있듯이 인공신경망도 이와 비슷한 구조를 가지고있다[18](Fig2-1).

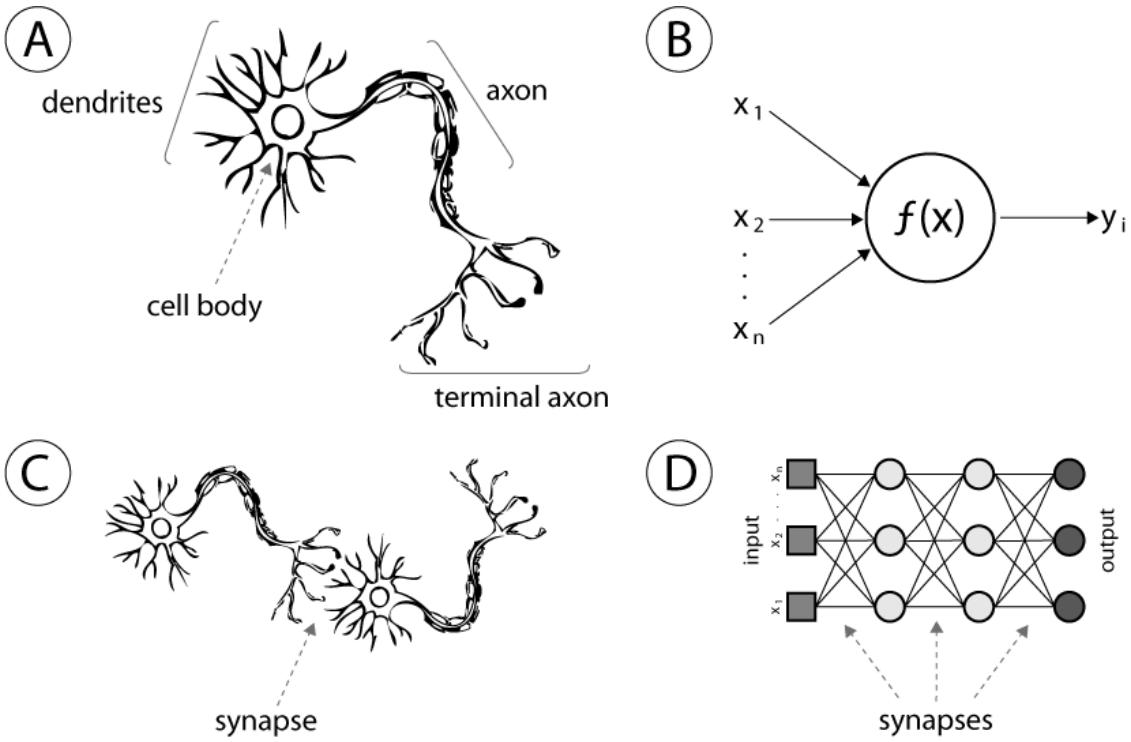

Fig 2-1 comparison between a human neuron and an ANN synapses[18)]

(A) human neuron; (B) artificial neuron; (C) biological synapse;

(D) artificial neural network synapses; ANN, artificial neural network

뉴런이 모이면 신경이 되듯이 인공 뉴런이 모이면 인공신경망이 된다. 신경망의 일반적인 알고리즘 구조는 입력층(input layer: 학습시키고자 하는 데이터)과 출력층(output layer: 학습된 모델에서의 결과) 사이에서 다수의 숨겨진 층(hidden layer)들로 이루어져 있다(Fig 2-2). 숨겨진 층은 앞선 입력층이나 다른 숨겨진 층의 가중치를 받아 다음 층으로 넘겨주는 역할을 한다. 입력층에서 데이터 정보를 넘겨주면 데이터의 각 특징(feature) 마다 숨겨진 층을 거치면서 가중치가 계산되고 이 값들은 최종적으로 출력층으로 이동하여 최종 값을 결정한다[19].

뉴런도 일정 이상 자극의 역치값을 넘어야 신호가 전달되듯이 여러 가지의 함수(function)의 역할로 임계선(threshold) 이상의 값이 도출되어야 다음 신경망으로 전달될 수 있다.

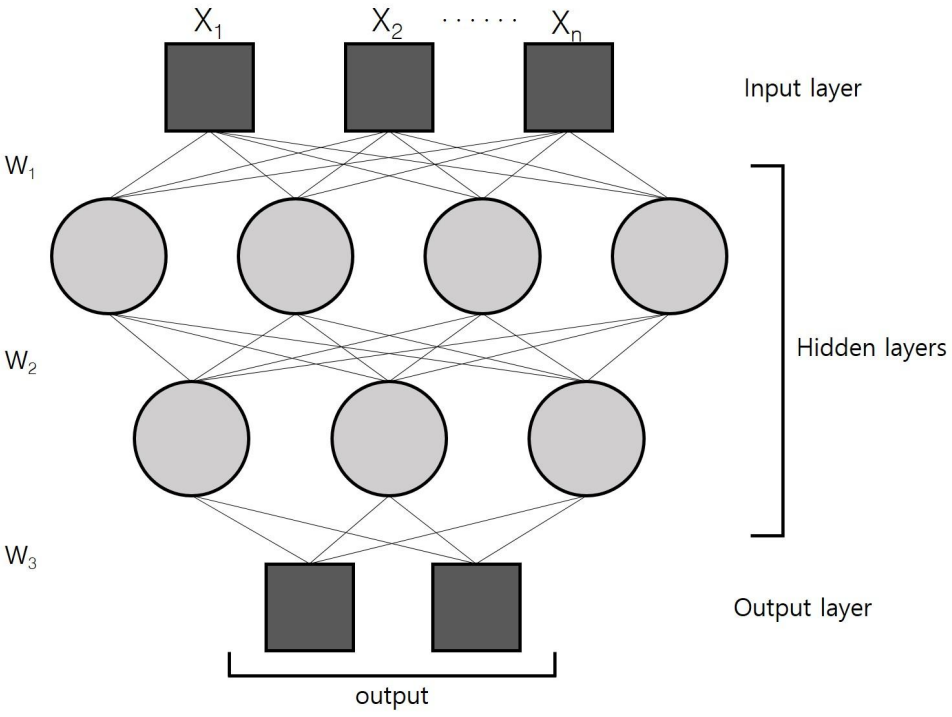

Fig. 2-2 An example of neural network.

4. 딥러닝(deep learning)

딥러닝은 머신러닝의 한 분야로 여러 비선형 변환기법을 조합하여 보다 데이터의 핵심적인 특징을 추출하고, 높은 수준의 추상화를 시도하는 머신러닝 알고리즘의 집합으로 정의된다[20].

딥러닝은 기본적으로 인공신경망의 구조를 기반으로 설계되어 있으나, 기존의 인공신경망의 단점인 과적합 현상, 느린 학습시간등을 보완하여 더욱 간단하고 성능이 좋은 형태로 발전하였다. 딥러닝을 구성할 수 있는 구조는 여러 가지가 존재하며 대부분의 경우 대표적인 몇 가지 구조들에서 파생된 것이며, 분석하고자 하는 데이터의 형태에 따라 알고리즘의 구조가 변할 수 있다. 대표적으로 영상 및 음성분야에서 좋은 성능을 보여주는 CNN과 필기체 인식과 같은 문자인식에 좋은 성능을 보여주는 RNN이 있다.

제 3 절 Convolutional Neural Network

1. Convolutional neural network의 구조와 이해

CNN은 영상 처리에 특화된 모델로서 의료영상분야에서 주로 채택하여 사용하고 있으며, 특히 당뇨성 망막증과 피부암을 탐지 및 분류하는데 있어서 CNN 모델은 높은 신뢰도와 효용성을 보여주고 있다[21,22]. 영상을 처리하는 방식에 있어서 기존의 신경망과는 다른 모습을 보여주는데 단순연산으로 영상의 특징을 추출하는 것이 아닌 특정한 필터(filter)를 이용하여 보다 복잡한 연산을 하도록 만들어 주는 것이 특징이다. 우리가 보는 영상은 픽셀 구조로 구성되어 있으며 위치 정보인 가로와 세로 그리고 회색 영상이 아닌 컬러 영상이라면 컬러 정보(RGB: Red, Green, Blue)를 추가하여 3차원 데이터가 된다(Fig. 2-3). 이러한 3차원 정보의 특징을 잘 추출하기 위해서 CNN의 구조는 convolutional layer와 max pooling layer, 마지막으로 fully connected layer로 구성되어 있으며, 출력값의 오차를 역방향으로 전달하여 가중치를 업데이트하는 역전파(back propagation) 방법을 사용한다(Fig. 2-4).

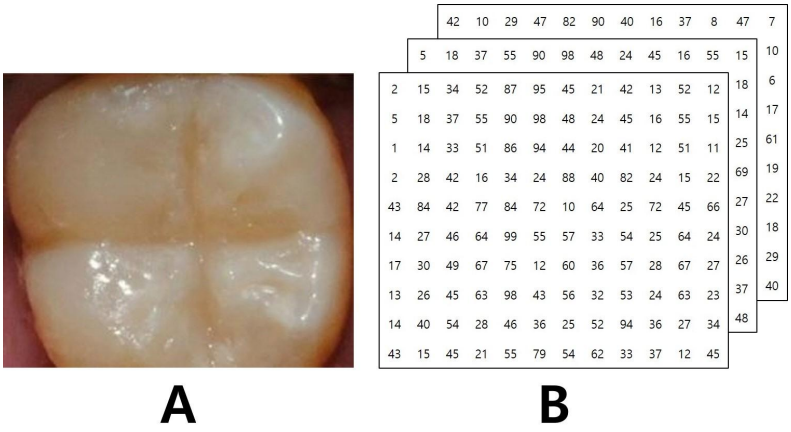

Fig. 2-3 Difference in image recognition between human and computer (A) What we see; (B) What computer see

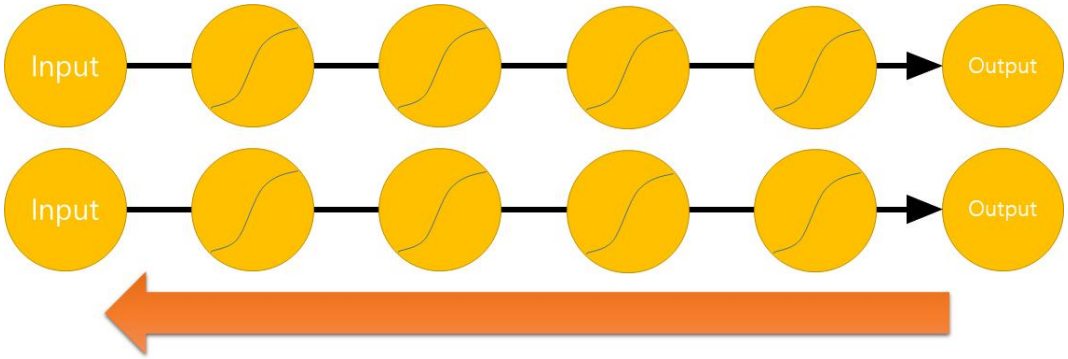

Fig. 2-4 Back propagation

1) Convolutional layer

3차원의 영상 데이터를 처리하기 위해서 특정한 사이즈의 커널(kernel)을 사용한다. 커널은 X축과 Y축을 곱한 2차원적인 개념으로 예를 들어 커널 사이즈(kernel size)를 3×3으로 정의를 내린다면, 이 커널은 각 칸마다 가중치(weight 혹은 parameter)가 존재하여 영상의 왼쪽 상단으로부터 한칸 씩 옆으로 이동하면서 원본 영상의 데이터 값(input)과 연산 과정을 거치게 된다. 이러한 커널 사이즈를 통한 연산 과정을 몇 번 거치게 되는지에 대한 정의가 필터 사이즈(filter size)이다. 필터는 Z축의 3차원 개념으로 커널 사이즈와 다른 개념이다(Fig. 2-5). 이러한 연산으로 얻어진 값들은 영상의 특성을 대표하며 feature라고도 말한다. 즉 Convolutional layer는 앞에서 전달받은 데이터의 정보를 커널 사이즈 및 필터 사이즈에 맞게 연산을 진행하여 다음 레이어(layer)로 값을 넘겨주게 되며 이러한 값들은 영상의 특성을 대표하게 되고, 영상을 잘 분류할 수 있도록 학습하게 된다.

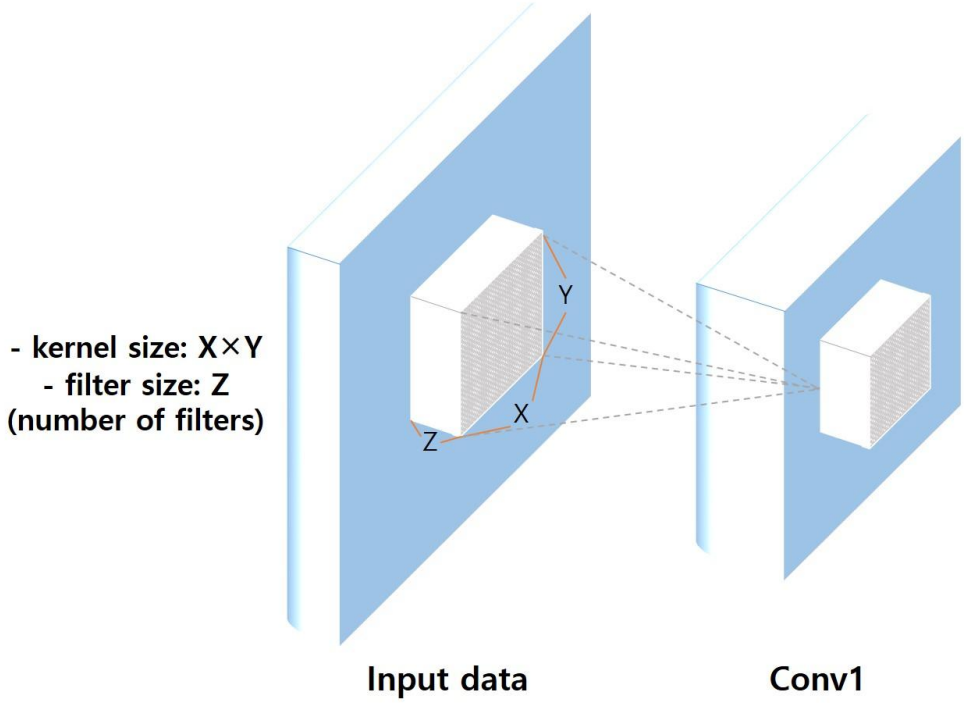

Fig. 2-5 Difference between kernel size and filter size

2) Pooling layer

풀링 레이어에서는 앞에서 전달받은 데이터 중 가장 의미있는 일부 데이터를 샘플링(sampling) 하여 데이터 정보량을 축소 시켜주는 역할을 한다[23]. 데이터를 샘플링 하는 방법으로는 최소값 풀링(minimum pooling), 평균값 풀링(average pooling), 최고값 풀링(max pooling)이 존재하며, 해당 영상의 특징을 가장 잘 반영할 수 있는 최고값 풀링을 일반적으로 사용한다.

Fig. 2-6 에서 하나의 4×4 size 데이터가 있다고 가정할 때 2×2 size의 filter size를 x축과 y축을 2칸씩 이동하면서 가장 큰 데이터 값만 뽑아낸다고 가정한다면 가장 특징적인 데이터를 추출하는 동시에 데이터의 양은 1/4로 줄일 수 있다.

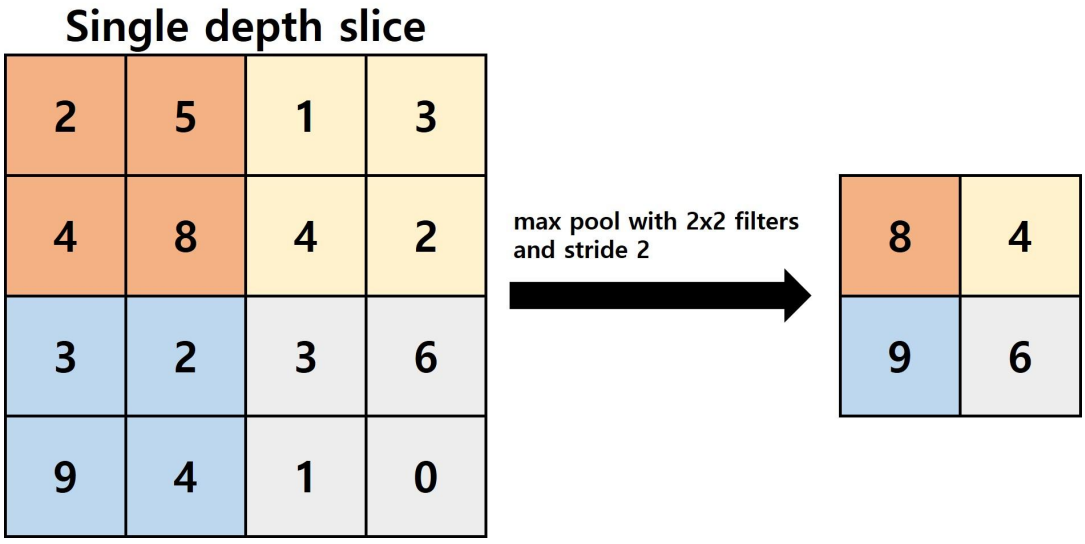

Fig. 2-6 An example of max pooling

3) Fully connected layer

Fully connected layer는 앞서 convolutional layer와 pooling layer를 통해 최종적으로 결과값을 나타내는 밀집되어 있는 층이다. 이 층에서 n차원의 벡터값을 갖는 확률을 나타내게 되는데 n의 개수는 우리가 확인하고자 하는 클래스(class)의 수와 동일하다. 1부터 10까지 10개의 숫자를 분류한다고 하면 최종 Fully connected layer는 10개의 결과 데이터를 생성해낸다. 이렇게 생성된 데이터는 확률로 변환하기 위하여 활성함수(activation function)를 사용하여 출력한다.

제 4 절 딥러닝 기술 용어

1. Cost function

비용함수(cost function)는 설계된 모델이 주어진 데이터에서 얼마나 잘 적용되어 오차가 적은지에 대하여 설명해 주는 함수이다[24]. 적은 값을 나타낼수록 더 적합한 모델이 설계되었다는 것을 의미한다. 회귀 모형에서도 설명변수에 의해 적합한 종속변수 값을 나타내기 위해 잔차(residual)제곱의 합을 최소화 하는 최소제곱법(method of least squares)이나, 잔차제곱의 평균을 나타내는 평균제곱오차(mean squared error) 방법을 사용하여 관찰된 자료에서 가장 잘 설명할 수 있는 직선을 구한다. 이와 같이 머신러닝에서도 비용함수의 값을 줄이기 위해 다양한 방법을 사용한다. 대표적인 예는 cross entropy이다. 주어진 집합의 분포 (p, q)에 대한 cross entropy는 다음과 같이 정의된다(Fig. 2-7).

$$CE = -\sum_{x} p(x) \log q(x)$$

Fig. 2-7 The formula of cross entropy

분포 p값에 대한 확률 q값은 0에서부터 1까지의 실수이므로 로그값을 취한다면 음수가 나오게 된다. 이 음수값을 보상하기 위에 로그값에 마이너스를 넣어 양의 실수가 나오도록 한다. 확률 q값이 1에 가까울수록 로그값은 0에 수렴하게 되므로 entropy(불확실성의 정도)가 낮아지게 된다. 확률 분포에서 표본을 뽑고, 이 분포와 목표 분포사이의 cross entropy를 최소화하여 다음

반복에서 더 적합한 가중치를 찾아가는 방법을 통해 비용함수를 줄여나간다.

2. Optimizer

optimizer는 오차에 대하여 가중치를 업데이트 시키는 알고리즘을 말한다. 가장 대표적으로 경사하강법(gradient descent)과 확률적 경사하강법(stochastic gradient descent, SGD)이 있다. 경사하강법은 가중치의 값에 따라 비용함수의 미분값 즉, 기울기가 양의 값이면 가중치를 감소시키고, 기울기가 음의 값이면 가중치를 증가시키는 방법이다(Fig. 2-8). 기울기의 값이 0에 수렴할수록 비용함수의 값 또한 감소하게 된다. 경사하강법의 단점은 수많은 단계마다 기울기를 구하는 방법으로 계산량이 방대해진다는 점이다. 이 단점을 보완하기 위해 SGD를 주로 사용한다. SGD는 매 단계마다 모든 데이터를 계산하지 않고 일부 데이터에서만 기울기를 계산하는 방식으로 같은 시간 안에 여러 번 기울기를 계산할 수 있으므로 빠른 속도로 가중치를 업데이트 할 수 있다.

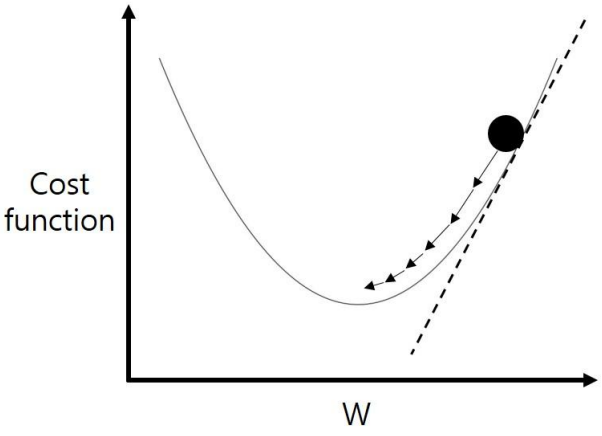

Fig. 2-8 Gradient descent

그러나 오차에 대한 전역 최소값(grobal minimum)이 있음에도 불구하고 최소 극소값(local minimum)에 머물러 더 이상 업데이트 되지 않는 문제가 발생할 수 있다(Fig. 2-9). 이러한 문제를 해결해주기 위해 momentum 기법을 사용한다.

Momentum의 사전적으로 '가속도'라는 의미를 갖고 있으며, 가중치를 업데이트하는 방향에 속도를 변화시켜 주는 방법으로서 현재 기울기에 방향으로 가속되어 업데이트된다. 즉, 오차 범위가 클 때 업데이트 간격을 더욱 벌려 최적값으로 가속되어 업데이트되므로 최소 극소값에 머무를 확률을 줄여준다.

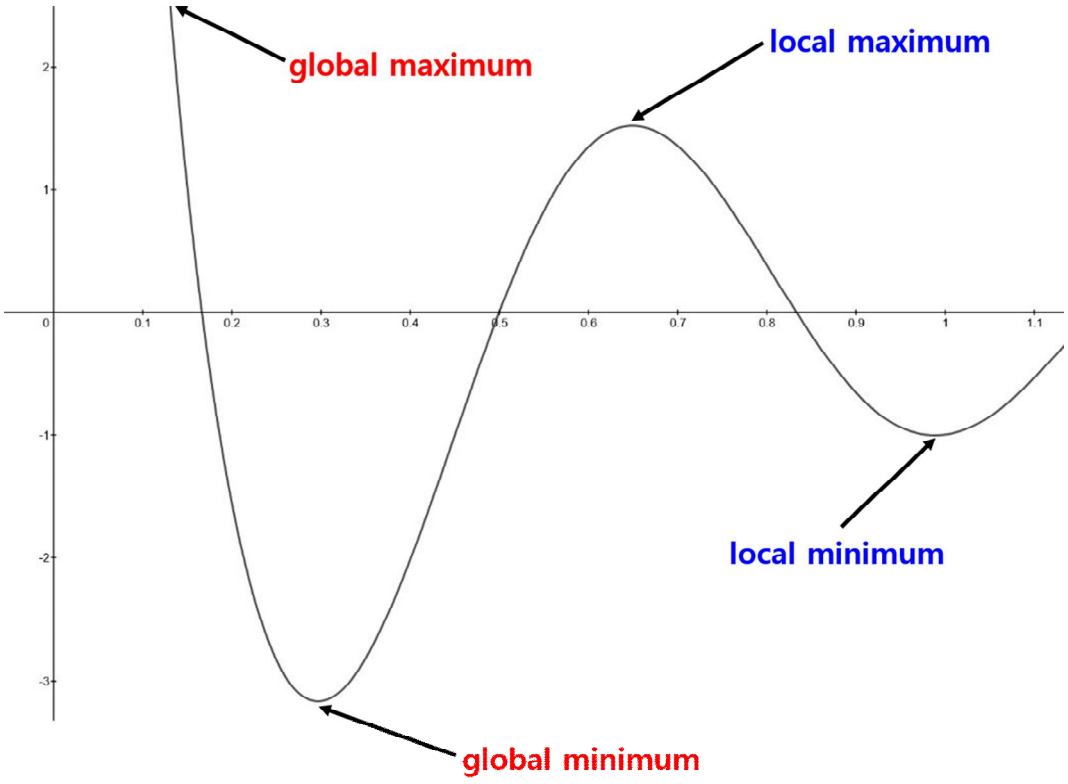

Fig. 2-9 Global minimum and local minimum

3. Activation function

convolutional layer를 거쳐 도출된 값들은 활성함수(activation function)를 사용하여 데이터의 특징을 더욱 부각시키고 정제시켜주는 역할을 한다. 상대적으로 불필요한 정보는 감소시켜주고, 비용함수를 줄여주는 핵심 입력 데이터에서 추출된 가중치는 증가시켜 준다. 대표적으로 3가지 활성 함수를 사용하여 다음 층으로 출력값을 내보낸다.

1) Sigmoid

데이터의 값을 0~1 사이 값으로 나타내고, 마이너스 입력값을 0에 가까운 값으로 만들어준다. 각 층의 말단 부위에서 이 함수의 영향을 받아 값의 형태가 변형되거나 정규화 된다(Fig. 2-10).

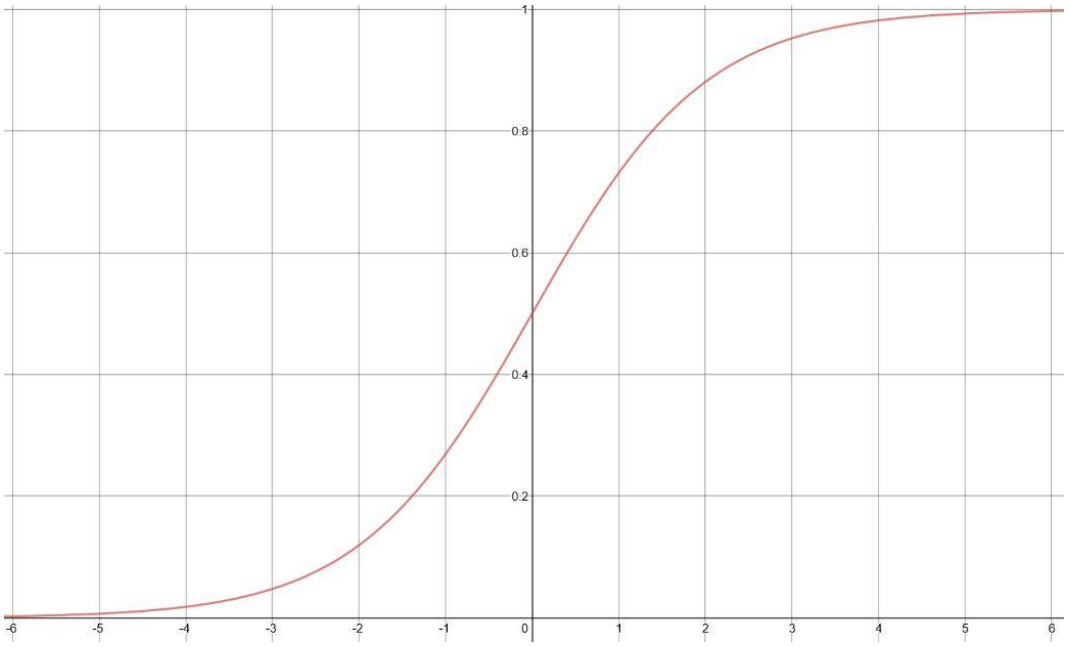

Fig. 2-10 Sigmoid function

2) ReLU

대표적으로 많이 사용하는 활성 함수로 음의 값을 갖는 값은 모두 0으로 만들어주는 특징이 있다[25]. 조금이라도 양의 값을 갖는다면 이 특성을 유지시켜주기위해 양의 기울기를 갖도록 연산한다. 역전파 방법을 사용하는 CNN모델에서는 sigmoid함수에 의해 모델의 신경망이 깊어질수록 기울기의 값이 사라지는 현상이 발생한다. 이를 기울기 소실의 문제(vanishing gradient)라고 하는데, 이를 효율적으로 방지해주는 것이 ReLU함수이다(Fig. 2-11).

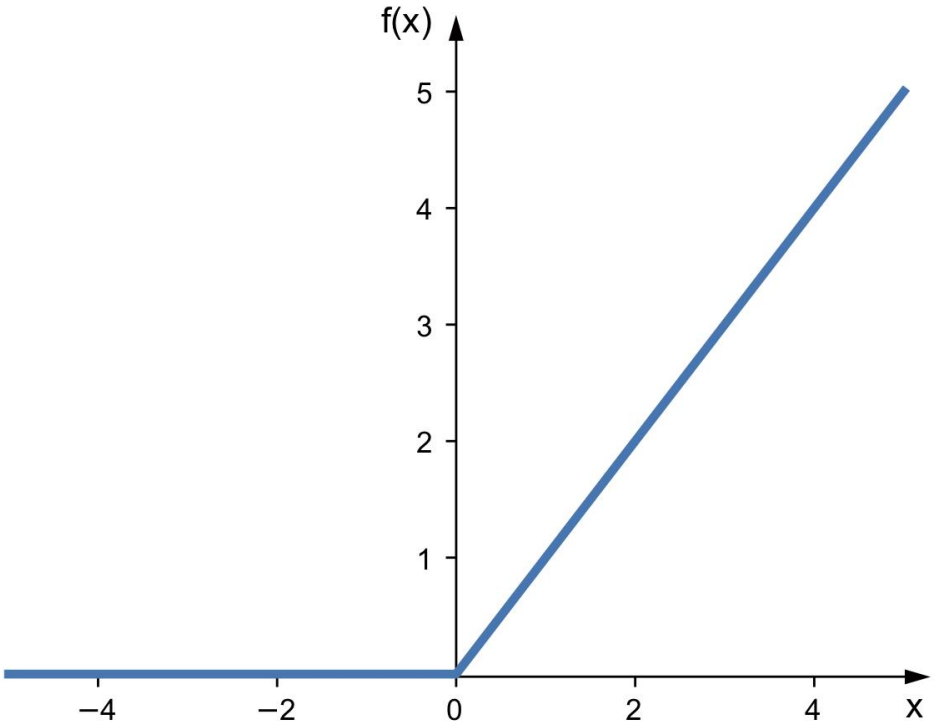

Fig. 2-11 ReLU function

3) Softmax

변수 n개에 대하여, 이 수의 대소 관계의 비율을 유지하면서 변수들의 합을 항상 1로 만들어주는 함수이다. 각 변수의 대한 확률값으로 변환하고자 할 때 사용되며 마지막 층에서 3개 이상의 범주로 나눌 때 주로 사용된다 (FIg. 2-12).

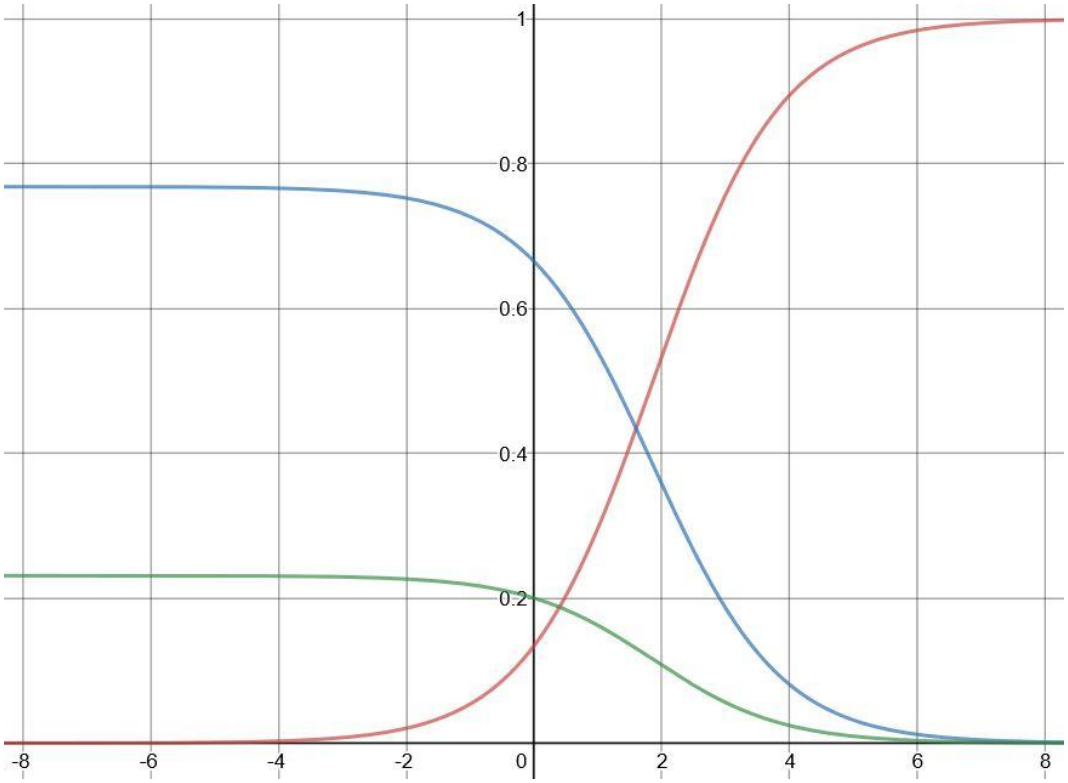

Fig. 2-12 Softmax function

제 5 절 프로그래밍 언어

1. 파이썬(python)

‘빅 데이터’(big data) 시대에서 대량의 데이터를 의미 있는 지식으로 변환하는 것은 다양한 분야에서 점점 더 중요해지고 있다[26]. 디지털 사회에서 수많은 데이터를 우리는 접하고 살아가고 있지만, 데이터를 효율적으로 관리하고 통계적인 방법으로 유의미한 지식으로 변환하는 것은 어렵다. 특히 정형화되어 있는 데이터를 처리하는데 있어, 비교적 접근성이 좋게 설계된 프로그램들이 존재하지만, 비정형 데이터를 처리하기 위해서는 데이터 포맷을 변환하지 않는 이상 바로 활용하기 힘들다. 이러한 비정형 데이터의 분석에 대한 접근성을 높여 주는 프로그래밍 언어는 파이썬, 지바, C 언어 등이 있으나 영상 분석에 가장 적합한 언어로는 파이썬을 대표적으로 사용한다.

파이썬은 플랫폼 독립적이며, 인터프리터식, 객체지향적, 동적 타이핑(dynamically typed) 대화형 언어이다[27]. 오픈소스 라이선스를 받아 안정성이 높으며 무료로 배포되고 있다. 원래 교육용으로 설계되었기 때문에 코드를 최대한 간단하게 쓸 수 있도록 하였으며, 가독성과 단순함을 강조한다. 이러한 강점으로 비정형 빅데이터를 처리하는 방법으로 파이썬 언어를 주로 사용하고 있으나, 단점으로는 들여쓰기 사용하여 블록을 구분하는 독특한 문법을 채용하고 있기 때문에 잠재적인 버그(bug)에 대한 디버깅(debuging)의 어려움이 있다라는 비평을 받기도 하였다[27].

2. 라이브러리(library)

파이썬 소프트웨어를 처음 설치할 때는 가장 기본적인 기능만을 사용할 수 있도록 설계되어 있으며, 이는 최소한의 기능만을 사용하면서 확장해 나갈 수 있도록 디자인되었다. 라이브러리는 파이썬에서 다양한 기능을 사용하기 위한 추가 구성 요소를 말한다. 딥러닝 분석을 하기 위해서 대중적으로 사용되는 몇 가지 라이브러리를 살펴보면 다음과 같다(Table 2-1).

Table 2-1. Python library for deep learning analysis

Library	Explanation
Tensorflow	Machine learning system
Keras	Designed to enable fast experimentation with deep neural networks.
Numpy	Adding support for large, multi-dimensional arrays and matrices
Matplotlib	A plotting library for the Python programming language and its numerical mathematics extension NumPy
tqdm	Check the progress of the command
Pandas	It offers data structures and operations for manipulating numerical tables and time series

제 3 장 연구대상 및 방법

제 1 절 연구 설계 및 대상

1. 연구 설계

이번 연구는 인터넷 검색 엔진에서 치아 영상을 수집하는 Search engine scraping 방법을 사용하여, 서울대학교 치의학대학원 연구윤리심의위원회의 승인 후에 진행하였다(No. S-D20190007). 2019년 4월부터 5월까지 약 2개월에 걸쳐 인터넷 검색 엔진 구글(Google)에서 특정 키워드(key word)에 노출되는 교합면이 보이는 치아 영상을 수집하였다. 선정기준과 제외기준에 부합되는 치아 영상을 선별하고 건전치아, 우식치아, 처치치아로 3가지 분류로 나누었다. 수집된 영상을 딥러닝 모델에 적합하도록 전처리 후에 훈련세트와 시험세트로 나누어 각 모델의 성능을 확인하였다. 최종 결과값으로 일치도, 민감도, 특이도, 양성예측도, 음성예측도, ROC(receiver operating characteristic) curve 및 AUC (area under the ROC curve)를 도출하여 모델의 신뢰도를 평가하였다. 연구의 흐름은 Fig. 3-1과 같다.

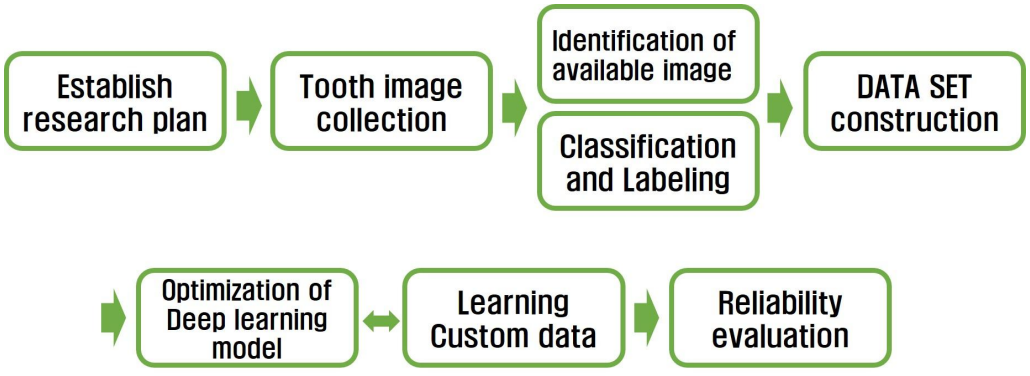

Fig 3-1. Research flow chart

2. 연구 대상

인터넷 검색 엔진 Google에서 특정 키워드를 검색하여 교합면이 보이는 소구치 및 대구치의 치아 영상을 수집하였다. 키워드는 'occlusal surface of sound teeth', 'occlusal surface of caries teeth' 'occlusal dental caries', 'dental restoration', 'tooth filling', 'dental crown'을 사용하여 검색하였다. 저작권 침해의 우려가 있는 사진이나, 개인식별이 가능한 사진, 치과 홍보 목적으로 업로드(upload)한 사진, 해상도가 낮아 분석하기 힘든 사진, 사진의 크기가 작아 분석하기 어려운 사진은 제외하였다. 또한 사진의 게시자가 다른 목적으로 활용하기를 원하지 않는다는 문구가 적힌 게시글은 수집하지 않았다. 수집된 영상의 분류는 WHO의 구강검진조사의 기준으로 건전치아, 우식치아, 처치치아로 나누었다. Table 3-1과 같은 선정기준과 제외기준에 부합하는 영상을 수집하였다.

Table 3-1. Inclusion/exclusion criteria for dental image

Inclusion criteria
1. 상하악 교합면이 보이는 소구치 및 대구치
2. 건전치아: WHO 구강 검진 기준에 부합하는 치아
3. 우식치아: 교합면에 와동이 형성된 치아, 명백한 법랑질 우식
4. 처치치아: 교합면에 수복물이 존재하는 치아(filling, inlay, onlay, crown)
Exclusion criteria
1. 상하악 전치부
2. 초기우식병소(white spot)
3. 저작권 침해의 우려가 있는 사진
4. 개인식별이 가능한 사진
5. 해상도가 낮은 사진 (resolution<150)
6. 매우 작은 크기의 사진 (pixel<100×100)
7. 치과 홍보용으로 업로드한 사진
8. 게시자가 다른 목적으로 활용을 원하지 않는 게시글의 사진

제 2 절 연구 방법

1. 영상 레이블링(labeling) 및 전처리 방법

수집된 영상은 명확하게 분류하기 위해 2명의 예방치과 전문의 치과의사가 판독하여 각 영상 상태에 맞는 정보를 기입하였다. 판독자간 일치도를 확인하기 위하여 1차로 210개의 치아 영상을 레이블링하여 kappa index를 구하였다(kappa index: 0.971). 일치하지 않는 4개의 영상은 제외하고 새롭게 일치된 치아 영상을 채워 최종적으로 각각 70개의 영상이 선택되었다.(건전치아: 70개, 우식치아: 70개, 처치치아: 70개)

총 210개의 원본 영상 A는 사이즈가 모두 달라 CNN 모델에서 학습의 용이성을 위해 224*224 픽셀(pixel) 크기로 조정 하여 모두 동일한 사이즈의 영상 B로 변환하였다(Fig. 3-2). 파일의 확장자(file name extension)또한 'JPG' 파일로 변환하였다. CNN 모델을 학습시키기 위한 훈련세트와 이를 검증하는 시험세트는 무작위로 각각 50개, 20개로 나누었다.

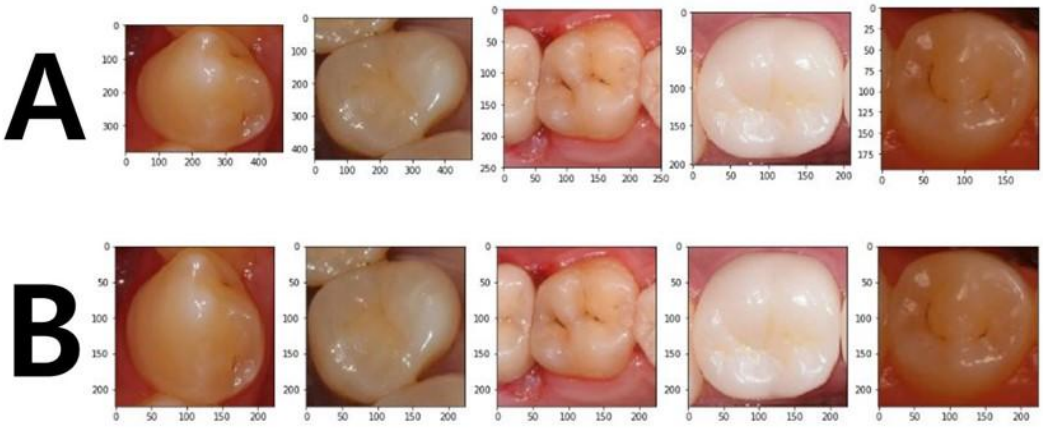

Fig. 3-2 Preprocessing: pixel size resizing
(A) image before pixel resize; (B) image after pixel resize

2. 통계분석

딥러닝 분석을 위한 모든 코드(code)는 파이썬 프로그래밍 언어를 기반으로 Tensorflow를 지원하는 Keras 라이브러리에서 작성되었다. 딥러닝 모델에서 건전치아와 우식치아 분류의 신뢰도를 평가하기 위해 각 치아 영상의 레이블링정보와 딥러닝 모델에서 fully connect layer에서 최종 산출된 output 값을 임계점으로 설정하여 ROC curve를 산출하였으며, 일치도, 민감도, 특이도, 양성예측도, 음성예측도, ROC curve의 내부 면적인 AUC를 산출하였다. 딥러닝 모델 구축을 위한 코드 작성 이외의 통계분석은 IBM SPSS Statistics 23.0 (IBM Co., Armonk, NY, USA)을 이용하여 분석하였다.

제 4 장 연구 결과

제 1 절 두가지 범주를 분류하는 모델

1. 전이학습 비적용 모델

먼저 CNN 모델을 구축할 때, 가지고 있는 영상의 픽셀량이 크지 않다고 판단하여 convolution layer 4개, max pooling layer 2개, fully connected layer 2개로 구성된 CNN 모델을 구축하였다. 총 8개의 층으로 이루어진 모델에 건전치아 영상과 우식치아 영상을 각각 50개를 훈련 세트에 넣어 학습시켰다. 비용함수를 최소화하기 위해 'binary cross entropy'를 활용하였으며, optimizer로서 SGD를 사용하였다. Learning rate는 10^{-6}의 값을 주어 미세한 변화에도 학습시키도록 하였으며 10회(epoch) 반복 학습하였다. 총 훈련된 parameter의 개수는 12,861,969개였다. 모델을 모형화하면 다음과 같다(Fig. 4-1).

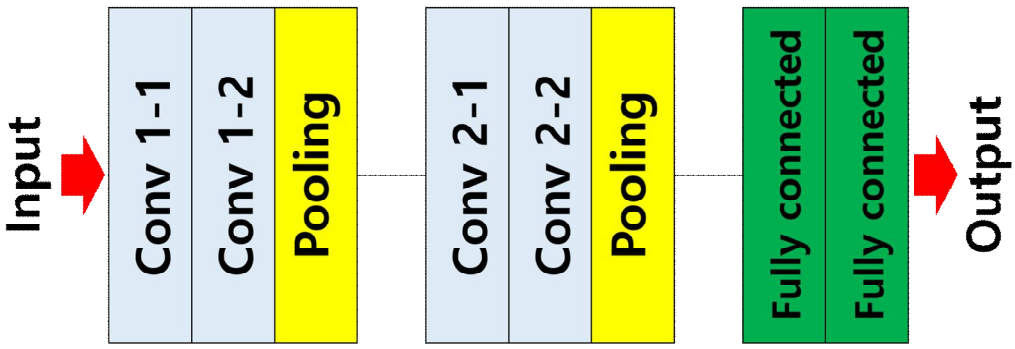

Fig. 4-1 Transfer learning non-use model

그 결과 손실값은 8.06, 일치도는 50%가 나왔다. 50%의 일치도는 두 개의 정답지 중 우연히 맞출 확률과 같으므로 전혀 학습이 되지 않았다라는 것을 의미한다. 이를 바탕으로 학습률을 높이기 위해 두가지의 결론을 얻었다. 첫 번째로 치과 영상의 특징을 추출하기 위해 더 깊은 신경망을 구축 해야한다. 두 번째로 특징 추출 층에 모델이 매번 새롭게 학습되기 때문에 전이학습을 적용한 CNN 모델 구축이 필요하다.

2. 전이학습 적용 모델

앞서 비교적 얕은 신경망으로 구축된 모델을 사용한 결과 손실값이 높게 나오며 학습률이 저조한 모습을 보여주었다. 이를 보완하기 위하여 2014년 영상인식 경진대회(ILSVRC)에서 준우승한 VGG16모델(Fig. 4-2)을 전이학습에 활용하여 치아 영상을 학습 후 예측결과를 확인하였다. VGG16은 convolution layer 13개, max pooling layer 5개, fully connected layer 3개로 구성된 모델이다. VGG모델은 가중치가 학습되지 않고 마지막 fully connect layer만 새롭게 학습되어져 학습량을 줄일 수 있다(Fig. 4-3). 전이학습에 활용된 VGG16의 총 parameter수는 17,926,209개이며 이 가중치를 활용하여 학습된 치아영상의 fully connect layer의 parameter 수는 3,211,521개이다. 손실값을 최소화 하기 위해 이전 모델에서 사용한 binary cross entropy와 optimizer로서 SGD를 활용한 것은 동일하나, local minimum에 빠져서 학습률이 저조해지는 것을 방지하기 위해 Momentum(0.9) 기법을 추가적으로 사용하였다.

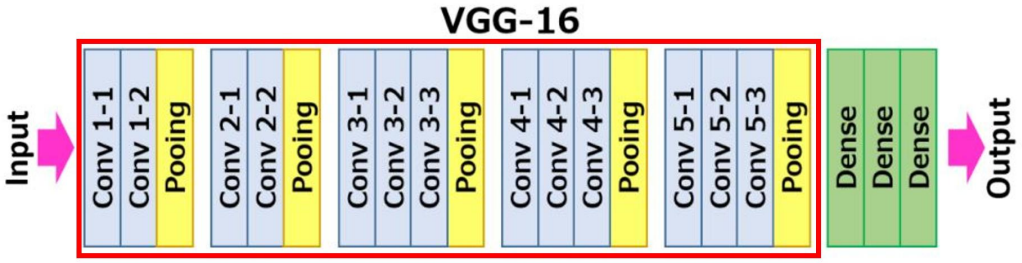

Fig. 4-2 VGG16 model: non-trainable parameter

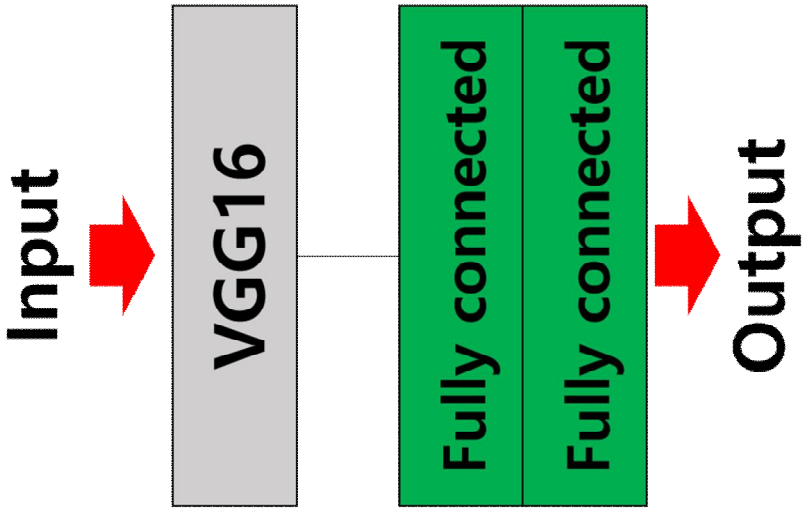

Fig. 4-3 Trainable parameter in transfer learning

10회(epoch=10) 반복 학습한 결과 손실값은 0.43, 일치도는 90%가 나왔다(Fig. 4-4). 작은 신경망을 구축한 모델에 비하여 확연히 일치도가 상승하였다. 학습 횟수(epoch)가 증가할수록 손실값의 감소하는 것으로 보아 치아 영상을 분류하는 알고리즘이 학습되고 있다는 것을 확인하였다.

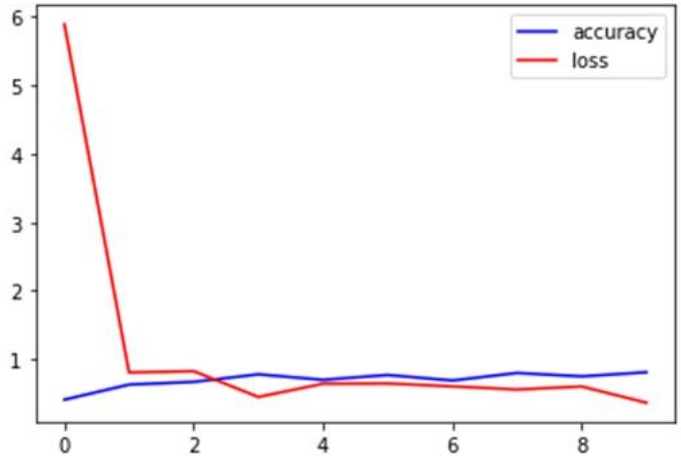

Fig. 4-4 Loss and accuracy by epoch in two classification model

CNN 모델을 활용하여 치아우식 탐지의 2*2 분할표는 아래와 같다(Table 4-1). 일치도, 민감도, 특이도, 양성예측도는 90%로 준수한 성능을 보였으며(Table 4-2), 이 결과는 학습 모델이 원하는 방향으로 제대로 학습이 되었다는 것을 확인할 수 있었다.

Table 4-1. Confusion matrix based on results of dental caries classification using CNN model

		Standard		
		Caries	Sound	Total
Predict	Caries	18	2	20
	Sound	2	18	20
	Total	20	20	40

CNN, convolutional neural network.

Table 4-2. Accuracy, precision, sensitivity, specificity, positive predictive value, and negative predictive value for the detection of dental caries

Accuracy	Precision	Sensitivity	Specificity	PPV	NPV
90%	90%	90%	90%	90%	10%

CNN, convolutional neural network; PPV, positive predictive value; NPV, negative predictive value.

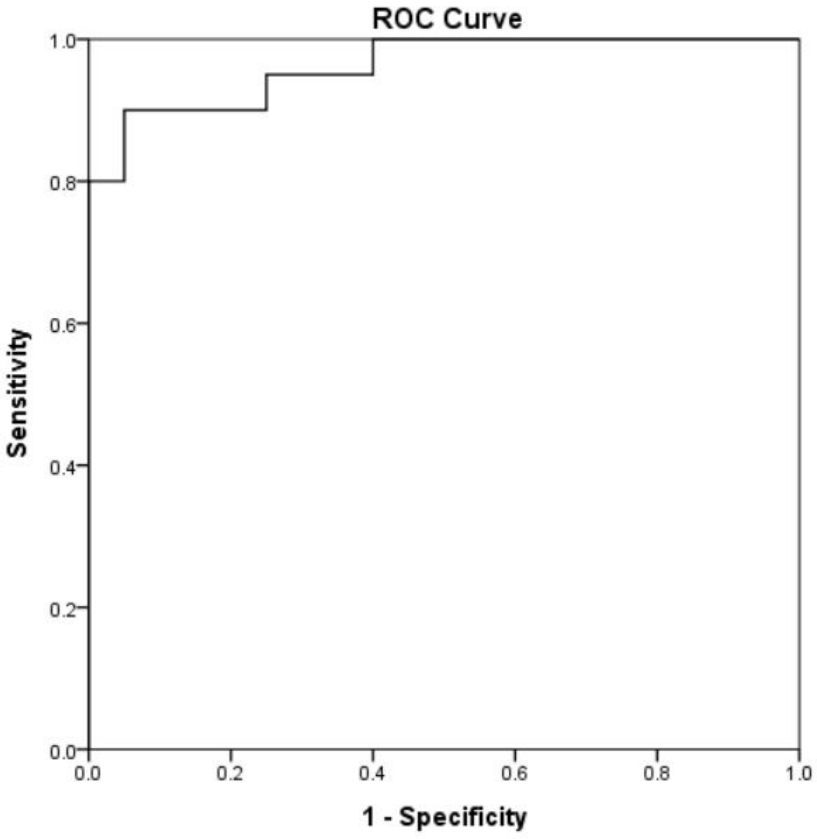

Fig. 4-5 Receiver operating characteristic curve

ROC curve의 내부면적인 AUC값은 0.963으로 높은 성능을 보여주는 것으로 확인되었다(Fig. 4-5). 비록 적은 숫자의 영상을 사용하여 학습시켰음에도 불구하고 전이학습을 적용한 모델에서 일치도가 크게 향상되었다.

제 2 절 세가지 범주를 분류하는 모델

1. 전이학습 비적용 모델

이미 두가지 범주를 분류하는 전이학습 비적용 모델에서는 학습 효과를 보이지 못하였으나 전이학습 전후의 학습 차이를 확인하기 위해 이전과 같은 모델을 사용하였다. 여기서 보철치료를 받은 처치치아를 포함하여 세가지 범주로 나누어 분류하고 그 결과를 확인하고자 하였다. 세가지 범주를 분류하는 모델에서도 마찬가지로 convolution layer 4개, max pooling layer 2개, fully connected layer 2개로 구성된 CNN 모델을 구축하였다. 건전치아, 우식치아, 처치치아 영상을 각각 50개를 훈련 세트에 넣어 학습시켰다. 비용함수를 최소화하기 위해 'categorical cross entropy'를 활용하였으며, optimizer, learning rate 등 모델 구축에 필요한 함수 또한 동일하게 설정하였다. 10회(epoch) 반복 학습하였다. 총 훈련된 parameter의 개수는 12,861,969개였다.

그결과 손실값은 8.26, 일치도는 33.3%가 나왔다. 33.3%의 일치도는 세 개의 정답지 중 우연히 맞출 확률과 같으므로 두가지 범주를 분류하는 모델과 같이 전혀 학습이 되지 않았다라는 것을 의미한다. CNN 모델이 치과 영상을 명확하게 학습하기 위해서는 전이학습의 적용이 필요하다고 판단하였다.

2. 전이학습 적용 모델

전이학습을 적용한 CNN 모델의 성능은 뛰어났으나, 건전치아와 우식치아의 분류는 이분형으로 결과가 나오는 모델이기 때문에 분류해야 할 범주의 개수가 많은 모델에 비하여 쉽게 답을 찾아냈다고 판단하였다. 두 번째와 같은 전이학습을 적용한 모델에서 적은 수의 영상을 보완하고자 데이터 증강(data augmentation) 방법을 사용하였다. 또한 학습 횟수(epoch)를 100회로 늘려서

손실값과 일치도의 추이를 확인하였다. 증강된 영상으로 학습을 마친 모델의 총 parameter수는 25,694,211이며 전이학습의 parameter까지 더한다면 총 40,408,899개의 parameter 수를 가졌다.

100회(epoch)의 학습이 끝난 결과, 약 50회(epoch)까지 손실값이 떨어지다가 그 이후로는 미미하게 하락하기는 하나 그 폭이 크지 않았다(Fig. 4-6). 일치도 또한 완만하게 상승하는 추이를 보였지만 후반부에서는 상승 폭은 거의 없었다. 임상 영상의 치아 상태에 따른 CNN 모델의 일치도는 다음의 표와 같다(Table 4-3).

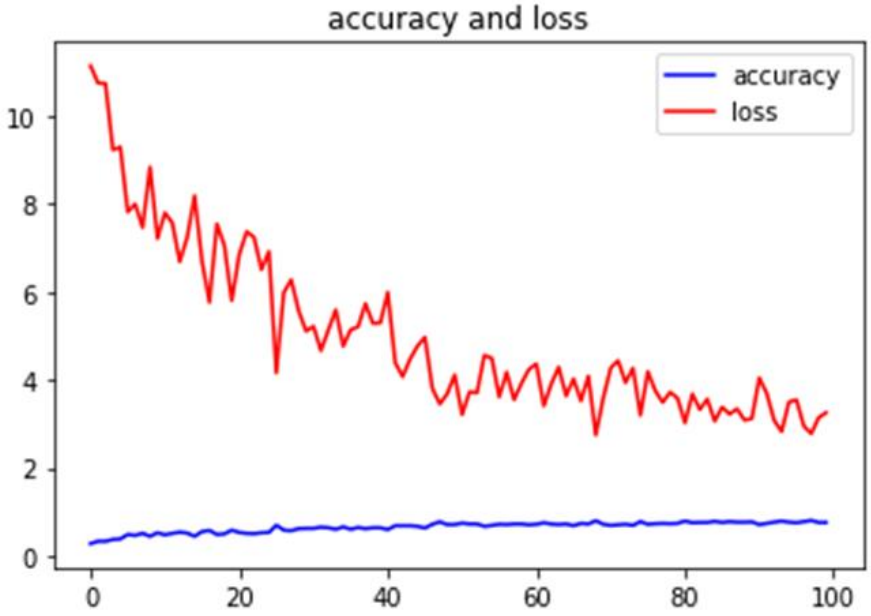

Fig. 4-6 Accuracy and loss by epoch in three classification model

Table 4-3. Accuracy of CNN model according to tooth status of clinical image

	Sound	Caries	Treatment	Total
Accuracy rate	25%	35%	80%	61.7%

CNN, convolutional neural network

건전치와 우식치아에서 매우 저조한 성능을 보여주고 있는 반면 처치치아에서는 준수한 성능이 나왔으나 전체적으로는 모델의 일치도는 많이 떨어진다. learning rate와 epoch 수를 조정하였으나 결과는 크게 다르지 않았다.

제 3 절 전이학습 적용 모델과 비적용 모델 비교

두가지 범주를 분류하는 모델과 세가지 범주를 분류하는 모델에서 모두 전이학습을 적용한 모델이 더 높은 성능을 보여주었다(Fig. 4-7), (Fig. 4-8). 치과 영상의 특징을 효과적으로 추출하고 신뢰도 높은 모델 구축하기 위해서는 전이학습의 적용이 필요하다는 결과를 얻었다.

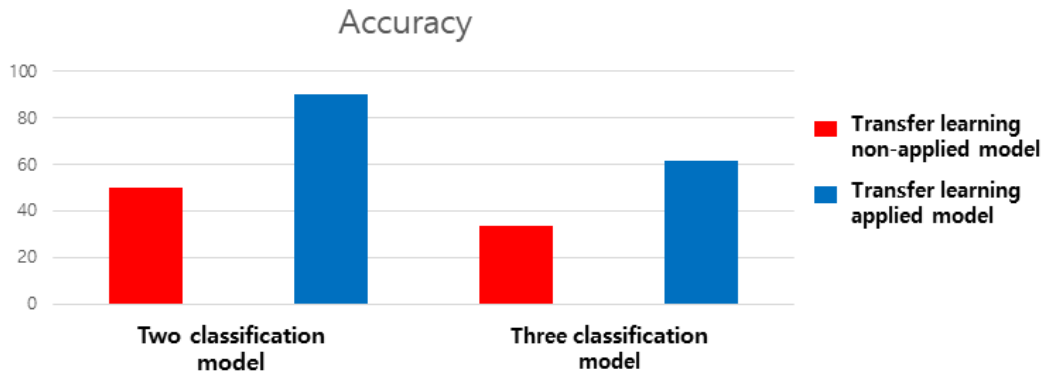

Fig. 4-7 Comparison of accuracy according to application of transfer learning

Fig. 4-8 Comparison of loss value according to application of transfer learning

제 5 장 고 안

인공지능의 기술이 나날이 발전해 가면서 공학분야를 넘어, 의료분야 및 치과분야까지 그 범위가 점점 넓혀지고 있다. 최근 치과 방사선 영상으로 치아우식을 탐지 및 진단하는 연구가 보고되었으나, 임상 치아 영상을 이용하여 탐지 및 진단하는 연구는 아직 보고 되지 않았다. 이번 연구에서는 컬러 치아 사진(RGB)을 이용하여 딥러닝 분석방법 모델로 치아우식 탐지의 신뢰도를 평가하였다.

이번 연구에서 전이학습 비적용 모델에서는 치아 영상에 대한 학습률이 저조하였으며, 일치도 또한 미미하였다. 치아 영상에 대한 학습률을 높이고자 이미 학습이 잘 이루어진 모델을 사용하는 전이학습 적용하여 우식치아와 건전치아의 분류의 일치도를 90%까지 올렸다. 그러나 처치치아를 포함하여 세가지의 치아 상태를 분류하는 모델에서는 일치도가 61.7%로 저조한 성능을 보여주었다.

치과 방사선 영상을 이용하여 치아우식을 탐지하는 선행 연구에서는 일치도가 약 82%로 보고되었다[2]. 낮은 일치도를 보인 것은 아니지만 임상에서 진단도구의 활용성 측면으로 보았을 때 신뢰할 수 있는 높은 수준의 결과는 아직 미흡한점이 있다고 판단하였다.

그러나 치과 분야에서 인공지능은 아직 초기 단계임에 불구하고 인공지능 응용 기술은 현저하게 발전하고 있다[28]. 영상 치의학 분야 이외에도 인공 신경망을 활용하여 치과 통증을 예측하는 연구가 보고 되었으며[29], 국민건강영양조사(NHANES) 자료의 사회경제적, 영양 정보를 활용하여 영구치우식 경험자를 분류하려는 시도가 있었다[30]. 또 다른 선행 연구에서는 다량의 전자 의무기록의 데이터를 활용하여 발치의 직접적으로 영향을 미치는 요소를 파악하고

자 하였다.

이러한 인공지능을 활용한 치과 질병 예측 모델은 의미있는 정보를 추출하여 치과 진단과 환자 자신의 질병을 이해하도록 돕는 측면에서 중요한 역할을 할 수 있다[16].

치과분야에서의 인공지능을 활용한 선행연구와 현황을 비추어 보았을 때, 이번 연구에서 CNN 모델을 활용한 우식치아와 건전치아의 성공적인 분류는 치과분야에서 임상 영상을 이용하여 판독하는것에 도움을 줄 수 있다는 관점에서 의미가 있다. 그러나, 수복물이 존재하는 치아를 포함하여 분류하는 모델에서는 성능이 급격히 떨어졌다. 이 결과에 있어 몇 가지 한계점이 존재한다.

첫 번째로 작은 데이터의 수이다. 딥러닝은 통상적으로 '빅데이터'를 이용하여 유의미한 정보를 추출한다는 의미에 주목하고 있다. 이번 연구에서 사용된 치아 영상의 수는 총 훈련세트 150개, 시험세트 60개의 작은 수로 충분히 학습이 되지 않았다라는 점을 간과할 수 없다. 또한 WHO의 구강검진기준을 따라 치아 영상을 분류하였으나 실제 임상에서 얻는 영상은 초기우식병소(white spot)이나 2차 우식증, 치간 우식등 숙달된 검진자가 여러 검진도구를 활용하여 진단을 내리는 치아상태일 경우 딥러닝 모델의 신뢰도는 더욱 떨어질것으로 예상되나, 각 영상의 특징을 추출할 수 있는 대량의 영상이 준비되어 있다면 좋은 성능을 낼 것이라고 사료된다.

두 번째로 치아 영상의 명확한 분류기준이 될 수 있는 조직학적 정보(histological score)가 없다. 이번 연구에서는 인터넷 검색 엔진에서 얻은 교합면이 보이는 치아 영상을 수집하였다. 데이터 수집 방법에 대한 한계점으로 치아 영상에 대한 조직학적 정보 및 병소의 심각도를 확인할 수 없으며, 빛의 양, 영상의 각도, 착색된 치아 등 오판할 수 있는 여러 가지의 편향들을 배제하지 못하였다. 조직학적 정보가 부족하게 되어 영상의 분류가 잘못되었을 경우 학습과정에서 가중치의 업데이트가 원활하게 이루어지기가 힘들며, 혼란을

야기할 수 있다.

그럼에도 불구하고 이번 연구 결과에서 확인할 수 있는 장점은 치과 영상의 활용 가능성이다. 디지털 공제 방사선 촬영술(digital subtraction radiography), 광 간섭성 단층촬영기술(optical coherece tomography) 전산화 단층 촬영술(computed tomography), 정량 형광 검사 방법(Quantitative light-induced fluorescence) 등과 같이 치아우식을 탐지할 수 있는 다양한 진단 도구에 비해 비교적으로 데이터 수집이 용이하고, 접근성이 우수하다.

이번 연구의 기대효과로서 국가적 조사(national survey)와 같은 대규모의 구강검진조사에 구강내 사진촬영의 과정이 추가되어 데이터 축적이 된다면, 검진 신뢰도 향상에 도움을 줄 것으로 사료된다. 구강 스캐너 혹은 고화질의 표준화된 사진을 수집한다면 딥러닝 모델 구축에 있어서 필요한 데이터 수집에 큰 도움을 줄 것이다. 향후 연구에 의해 신뢰도 높은 모델이 구축된다면, 검진자의 진단에 의한 레이블링(labeling) 작업 및 전산화 작업 또한 막대한 노동력과 시간이 소요되므로 이 작업에 대한 보조 도구로서 사용할 수 있을 것으로 사료된다.

이번 연구는 딥러닝 분석 방법을 적용한 모델을 활용한 치아우식 탐지의 신뢰도를 평가로서 분석 방법에 대한 높은 이해도가 요구된다. 여러 가지 딥러닝 모델은 대부분 오픈 소스(open source)로 누구나 접근하여 사용할 수 있도록 공개 되었지만 분석을 실행하는 인터페이스(interface)는 컴퓨터 언어의 문법(syntax)기반의 프로그래밍 언어이다. 복잡한 명령어 기반으로 작동하는 프로그래밍 언어에 익숙하지 않은 연구자 입장에서는 사용자 친화적(user friendly)인 관점에서 접근성이 떨어진다는 단점이 있다. 그럼에도 불구하고 치과분야에서 딥러닝 분석을 진단의 정확성을 높일 수 있는 도구로서 활용된다면 치과병·의원에서도 실효성이 높을 것으로 사료된다.

이번 연구는 적은 수의 영상을 활용하였음에도 불구하고 좋은 성능을 보여

주었다는 점에서 향후 치과분야 딥러닝 분석 방법의 활용도는 더욱 확대될 것으로 기대된다.

제 6 장 결 론

제 1 절 연구 결론

딥러닝 분석을 활용한 치아우식 탐지의 신뢰도를 평가하고자 교합면이 보이는 건전치아 70개, 우식치아 70개, 처치치아 70개, 총 210개의 영상을 수집하고, 각 영상을 훈련세트 50개, 시험세트 20개를 무작위로 할당하였다. 각 모델의 단점을 보완하고 딥러닝 모델을 구축하고 평가하였으며, 다음과 같은 결론을 얻었다.

1. 치과 영상의 특징을 추출하여 신뢰도 높은 모델을 구축하기 위해서는 깊은 신경망으로 설계되어야 한다.
2. 건전치아와 우식치아를 분류하는 모델에서, 전이학습 적용모델은 전이학습 비적용 모델(50.0%)에 비하여 월등히 높은 일치도를 나타내었다(90.0%, AUC:0.963).
3. 건전치아, 우식치아, 처치치아를 분류하는 모델에서 일치도는 감소하였고(61.7%), 각 영상의 일치도는 처치치아(80.0%), 우식치아(35.0%), 건전치아(25.0%) 순이었다.

제 2 절 향후 연구

이번 연구는 임상사진에서 딥러닝 분석 활용의 가능성을 확인하였으나, 제한점 또한 확인할 수 있었다. 향후 연구에서는 여러 가지 제한점을 보완하기 위해 상당한 수와 다양한 질병 상태의 치아 영상을 수집하고 분석할 수 있는 연구가 필요하다. 특히, 우식치아의 경우 단순히 판단기준을 이분법적 분류인 우식과 비우식으로 나눌 것이 아니라, 임상진단에 널리 활용될 수 있는 International Caries Detection and Assessment System (ICDAS)의 평가 기준으로 분류하는 연구가 필요하다[31].

아직까지 치과분야에서 딥러닝을 활용한 의료 영상 분석은 주로 방사선 영상에서의 질병 탐지 및 분류에 중점을 두었으나, 최근 치아와 치아의 구조를 세분화할 수 있는 객체 탐지(object detection) 분석방법을 활용한 연구가 보고되었다[32,33]. 임상 사진에서 객체 탐지 분석방법을 적용한다면 치아의 종류와 위치를 탐지하고 Federation Dentaire Internationale (FDI) 치아 번호 체계에 의해 세분화할 수 있는 연구가 가능하다. 실용적으로 치과 임상 영상을 활용하기 위해 치아의 상태와 위치 정보까지 자동화하여 데이터가 저장될 수 있는 연구가 필요하다.

향후에 임상 치과 사진을 활용한 딥러닝 연구가 지속되어 발전한다면 치과 방문에 제약이 큰 대상자, 치과 공포증이 있는 대상자, 자신이나 자녀의 치아 상태를 알아보고 싶은 대상자들을 위하여 간편하게 핸드폰으로 치아 사진을 찍어 상태를 확인할 수 있는 진단 기술 개발 또한 가능성이 있다.

References

1. Featherstone JD. The caries balance: the basis for caries management by risk assessment. Oral Health Prev Dent 2004;2 Suppl 1:259-264.
2. Lee JH, Kim DH, Jeong SN, Choi SH. Detection and diagnosis of dental caries using a deep learning-based convolutional neural network algorithm. J Dent 2018;77:106-111.
3. Kuchenbecker KJ, Parajon RC, Maggio MP. Evaluation of a Vibrotactile Simulator for Dental Caries Detection. Simul Healthc 2017;12(3):148-156.
4. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015;521(7553):436-444.
5. Siau K, Lacity M, Sauter V. Panel: Artificial Intelligence and The Future of Work. MWAIS 2018; Proceedings. 56.
6. 김한준. 4차 산업혁명과 직업세계 변화 전망. 상담과지도 2018;53(0):17-43.
7. Ekert T, Krois J, Meinhold L, Elhennawy K, Emara R, Golla T, et al. Deep Learning for the Radiographic Detection of Apical Lesions. J Endod 2019;45(7):917-22.e5.
8. Murata M, Ariji Y, Ohashi Y, Kawai T, Fukuda M, Funakoshi T, et al. Deep-learning classification using convolutional neural network for evaluation of maxillary sinusitis on panoramic radiography. Oral Radiol. 2018 Dec 11 [Epub]. DOI:10.1007/s11282-018-0363-7.
9. Kositbowornchai S, Plermkamon S, Tangkosol T. Performance of an artificial neural network for vertical root fracture detection: an ex vivo study. Dent Traumatol 2013;29(2):151-155.
10. McCorduck P. Machines who think: a personal inquiry into the history and prospects of artificial intelligence. 2nd ed. Natick:A.K. Peters.; 2004:5.
11. Russell SJ, Norvig P, Canny J. Artificial Intelligence: A Modern Approach. 2nd ed. New Jersey:Prentice Hall;2003:939.
12. Crevier D. AI: The Tumultuous Search for Artificial Intelligence. 1st ed. New York:Pearson.; 1993:46-51.
13. Kaplan A, Haenlein M. Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. Business Horizons 2019;62(1):15-25.

14. The University of EDINBURGH School of Informatics. History of AI at Edinburgh, Artificial Intelligence at Edinburgh University : a Perspective [Internet]. [cited 2019 May 02]. Available from: http://www.inf.ed.ac.uk/about/AIhistory.html.
15. The New York Times. IBM Is Counting on Its Bet on Watson, and Paying Big Money for It [Internet]. [cited 2019 May 02]. Available from: https://www.nytimes.com/2016/10/17/technology/ibm-is-counting-on-its-bet-on-watson-and-paying-big-money-for-it.html?emc-=edit_th_20161017&nl=todaysheadlines&nlid=62816440.
16. Park WJ, Park JB. History and application of artificial neural networks in dentistry. Eur J Dent 2018;12(4):594-601.
17. Mitchell TM. Machine Learning. 1st ed. New York:McGraw-Hill Education; 1997:2.
18. Maltarollo V, Honório K, da Silva A. Applications of artificial neural networks in chemical problems. In: Suzuki K, Artificial Neural Networks-Architectures and Applications. Rijeka: IntechOpen.; 2013:203-212.
19. Zell A. Simulation Neuronaler Netze [Simulation of Neural Networks] 1st ed. in German:Addison-Wesley; 1994:72.
20. Definition of Deep learning [Internet]. Wikipedia; [cited 2019 May 7] Available from: https://en.wikipedia.org/wiki/Deep_learning.
21. Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanswamy A, et al. Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs. Jama 2016;316(22):2402-2410.
22. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, et al. Dermatologist-level classification of skin cancer with deep neural networks. Nature 2017;542(7639):115-118.
23. Liu L, Shen C, Hengel AVD. Cross-Convolutional-Layer Pooling for Image Recognition. IEEE Trans Pattern Anal Mach Intell 2017;39(11):2305-2313.
24. Sun W, Poot DHJ, Smal I, Yang Xm Niessen WJ, Klein S. Stochastic optimization with randomized smoothing for image registration. Med Image Anal 2017;35:146-158.
25. 김철우. 대소니의 딥러닝 기초(파이썬, 케라스로 배우는): 1판. 서울:디지털북스; 2018;224-225.

26. Manyika J, Chul M, Brown B, Bughin J, Dobbs R, Roxburgh C, et al. Big data: the next frontier for innovation, competition, and productivity. Chicago: McKinsey & Company; 2011:23-26.
27. Python [Internet]. Wikipedia; [cited 2019 May 7] Available from: https://ko.wikipedia.org/wiki/%ED%8C%8C%EC%9D%B4%EC%8D%AC.
28. Khanna S. Artificial intelligence: contemporary applications and future compass. Int Dent J 2010;60(4):269-272.
29. Kim EY, Lim KO, Rhee HS. Predictive modeling of dental pain using neural network. Stud Health Technol Inform 2009;146:745-746.
30. Zanella-Calzada LA, Galvan-Tejada CE, Chavez-Lamas NM, Rivas-Gutierrez J, Magallanes-Quintanar R, Celaya-Padilla JM, et al. Deep Artificial Neural Networks for the Diagnostic of Caries Using Socioeconomic and Nutritional Features as Determinants: Data from NHANES 2013(-)2014. Bioengineering (Basel). 2018;5(2).
31. Gugnani N, Pandit IK, Srivastava N, Srivastava N, Gupta M, Sharma M. International Caries Detection and Assessment System (ICDAS): A New Concept. Int J Clin Pediatr Dent 2011;4(2):93-100.
32. Chen H, Zhang K, Lyu P, Li H, Zhang L, Wu J, et al. A deep learning approach to automatic teeth detection and numbering based on object detection in dental periapical films. Sci Rep 2019;9(1):3840.
33. Tuzoff DV, Tuzova LN, Bornstein MM, Krasnov AS, Kharchenko MA, Nikolenko SI, et al. Tooth detection and numbering in panoramic radiographs using convolutional neural networks. Dentomaxillofac Radiol 2019;48(4):20180051.

Abstract

Reliability evaluation of dental caries detection using deep learning

Kim Seon-Jip, BSDH, RDH

(Directed by Assistant Prof. Cho Hyun-Jae, DDS, PhD)

Department of Preventive Dentistry and Public Oral Health,

The Graduate School, Seoul National University

Dental explorer and dental radiographs are mainly used for the examination of dental caries, however they tend to depend on the experience of the examiner in the end and the reliability of the examiner may not be satisfactory. To overcome this problem, we tried to improve the diagnostic accuracy of dental caries by applying the deep learning which processes images or recognizes patterns and has superior performance in image analysis.

This study was approved by the Institutional Review Board of the Graduate School of Dentistry, Seoul National University using the search engine scraping method for collecting tooth images in the internet search engine. From April 2019 to May 2019 it has been going on for about two months. Tooth images were divided into sound teeth, caries teeth, and treated teeth according to WHO guidelines. Finally, 70 images were collected respectively, and 50 training sets for learning the deep learning model and 20 test sets for verifying the training set

were randomly assigned. ROC curve, AUC, accuracy, sensitivity, specificity, positive predictive value, and negative predictive value were calculated to confirm reliability in the classification model of sound teeth and caries teeth, and, the model for classifying sound teeth, caries teeth, treated teeth confirmed the accuracy of the classification.

The following conclusion was obtained from the experiment.

1. In order to extract the characteristics of the dental image and establish a reliable model, it should be designed as a deep neural network.
2. In the model dividing sound teeth and caries teeth, the transfer learning model showed a much higher accuracy (90.0%, AUC: 0.963) than the non-transfer learning model (50.0%).
3. Accuracy was reduced (61.7%) in the classification of sound teeth, caries teeth, and treated teeth. The accuracy of each image was 80.0% for treatment, 35.0% for carious teeth, 25.0% for sound teeth, respectively.

Keywords : artificial intelligence, convolutional neural network, deep learning, dental caries

Student Number : 2017-29616