

碩士學位論文

AD-ResUnet: 심층 신경망을 이용한
치과 파노라마 X-선 영상
향상 기술 개발

嘉泉大學校 大學院

融合醫科學科

醫用生體工學 專攻

姜 熙 根

碩士學位論文

AD-ResUnet: 심층 신경망을 이용한
치과 파노라마 X-선 영상
향상 기술 개발

AD-ResUnet: Development of Panoramic Dental X-ray Image
Improving System using Deep Learning Network

嘉泉大學校 大學院

融合醫科學科

醫用生體工學 專攻

姜 熙 根

碩士學位論文
指導教授 孫暎敦

AD-ResUnet: 심층 신경망을 이용한
치과 파노라마 X-선 영상
향상 기술 개발

AD-ResUnet: Development of Panoramic Dental X-ray Image
Improving System using Deep Learning Network

위 論文을 融合醫科學 碩士學位 論文으로 提出함.

2020 年 12 月

嘉泉大學校 大學院

融合醫科學科

醫用生體工學 專攻

姜 熙 根

이 論文을 姜 熙 根의
工學碩士 學位論文으로 認准함

2020 年 12 月

審査委員長	조 장 희	㉠
-------	-------	---

審 査 委 員	손 영 돈	㉠
---------	-------	---

審 査 委 員	김 행 근	㉠
---------	-------	---

국문 초록

파노라마 방사선 촬영기(panoramic radiograph)는 대부분의 치과진료에서 진단 과정에 있어 가장 기본적이고 중요한 도구로 사용되고 있다. 하지만 파노라마 치과 X-선 영상(panoramic dental X-ray image)은 실제 치아 구조에 비해 왜곡이 크다는 단점을 갖고 있고, 또한 경추와 연조직과 같은 치과 진단에 방해가 되는 불필요한 신체부위들을 포함하고 있다. 이를 해결하기 위해, 전산화 단층 촬영(Computed Tomography; CT) 영상을 통해 왜곡이 적은 시뮬레이션 파노라마 영상을 생성하는 자동화 시스템 구축과 관련된 많은 연구들이 이루어졌다. 하지만 이는 결국 실제 파노라마 영상을 향상시키는 것이 아닌, CT 영상에만 의존하여 새로운 시뮬레이션 영상을 생성한다는 것이다.

본 연구에서는 이와 같은 문제들을 해결하기 위해서 실제 파노라마 영상을 향상된 파노라마 영상으로 재생성해주는 심층 신경망 기반의 AD-ResUnet(Asymmetric Deep-ResUnet)을 제안하였다. 이는 영상 생성 네트워크로 널리 사용되는 인코더-디코더(encoder-decoder) 구조를 사용하였고, 네트워크의 모든 층은 사전 활성화 함수 잔차 블록(Pre-activation Residual block) 구조를 적용하였다. 해당 네트워크를 훈련함에 있어 MSE(Mean squared error)와 SSIM(Structural Similarity Index Map), 두 개의 목적함수를 사용하여, 실제 파노라마 영상의 불필요한 요소들을 잘 제거할 수 있도록 하였다. 또한 최적의 훈련을 위하여, 하이퍼 파라미터 튜닝(hyper-parameter tuning)을 통해 훈련에 사용한 두 개의 목적함수 비율을 조정하였고, 선정된 하이퍼 파라미터를 통

해 제안한 네트워크의 최종 훈련을 진행하였다.

제안한 네트워크와 동일한 인코더-디코더 구조를 갖는 기존의 생성기(generator) 모델들과 본 연구에서 제안한 네트워크 모델을 통해 실험 데이터를 처리한 결과를 비교하여, AD-ResUnet의 성능을 평가하였다. 비교 지표로는 MSE, PSNR(Peak Signal-to-Noise Ratio), SSIM을 사용하였고, 그 결과, AD-ResUnet은 각각 0.00069, 31.775331 0.94658로 비교 대상 네트워크들에 비해 영상 향상 능력에 있어 뛰어난 성능을 보였다.

핵심 용어: 전산화 단층촬영 영상, 파노라마 치과 X-선 영상, 심층 신경망, 영상 향상 기법

목차

국문초록	5
목차	7
표 목차	9
그림 목차	10
I. 서론	11
II. 연구 방법	13
1. 데이터	13
2. 파노라마 시뮬레이션 영상 생성 시스템	13
3. 제안하는 파노라마 영상 향상 네트워크	17
1) 관련 연구	17
(1) Residual Unit	17
(2) Deep Residual Unet	18
2) AD-ResUnet : 파노라마 영상 향상 네트워크	19
4. 목적 함수	23
5. 훈련 방법	24
6. 평가 지표	26
III. 연구 결과	28
1. 하이퍼 파라미터 튜닝 결과	28
2. 성능 비교	29
3. 실제 파노라마 영상에 대한 네트워크 성능	34
IV. 결론	36
V. 참고 문헌	38

VI. Abstract	41
감사의 글	43

표 목차

표 2.1 제안한 AD-ResUnet의 정보	22
표 2.2 사용한 하드웨어 사양	25
표 3.1 Hyper-parameter tuning 결과 (AD-ResUnet)	28
표 3.2 결과 영상들에 따른 네트워크 별 성능 비교	30
표 3.3 결과 영상들의 관심 영역에 따른 네트워크 별 성능 비교 ...	30

그림 목차

<그림 2.2.1> Panoramic Image Generator의 출력 영상	16
<그림 2.3.1> Pre-activation residual unit 구조	18
<그림 2.3.2> Encoder-Decoder 구조와 U-Net 구조 비교	19
<그림 2.3.3> 본 연구에서 제안한 AD-ResUnet 구조	20
<그림 3.2.1> 실험용 영상 1 에 대한 네트워크 출력 영상	31
<그림 3.2.2> 실험용 영상 2 에 대한 네트워크 출력 영상	32
<그림 3.2.3> 실험용 영상 3 에 대한 네트워크 출력 영상	33
<그림 3.3.1> AD-ResUnet을 통해 처리한 실험용 영상	35

I. 서론

인공지능을 이용한 기술들이 활발히 연구됨에 따라 인공지능 기술을 의료분야에 적용시키려는 연구들이 활발히 진행되고 있다. 이 중 심층 신경망(Deep neural network)을 의료 영상처리 기술, 의료 데이터 분석, 의료 진단과 같은 의료 전반에 걸쳐 적용되는 등, 매우 광범위하고 빠르게 확장되고 있다. 이에 따라 의료 영상관련 연구자들은 흉부 X-선 영상을 기반으로 한 자동 진단 시스템 개발[1], 치주 골 손실 자동 검출[2]등의 의료 영상 처리 분야에 심층 신경망을 적용한 연구들을 활발히 진행하고 있다.

의료영상들 중 파노라마 방사선 촬영(Panoramic Radiograph)영상은 대부분의 치과 진료에서 골격 구조 평가를 위한 중요한 도구로 사용되고 있다. 최근까지 디지털 방사선 영상화 기법 등의 방사선 촬영 기술의 발전으로 저 선량의 방사선으로도 상당히 정확한 영상을 제공한다[3]. 하지만 방사선 촬영 기법이 갖는 왜곡 문제는 파노라마 방사선 영상이 갖는 알려진 한계 중 하나이고[[4], [5]], 또한 방사선 영상에는 경추와 연조직과 같은 진단에 방해가 되는 불필요한 신체부위들이 포함되어 있다. 이 때문에 보다 적절한 진단과 치료 계획을 위해서는 파노라마 방사선 영상 이외의 추가적인 전산화 단층 촬영 영상(Computed Tomography; CT)이 요구된다[6].

실제로 대부분의 CT 스캐너들은 CT 영상을 통해 실제 파노라마 X-선 촬영 영상보다 왜곡이 적은 가상의 시뮬레이션 파노라마 영상을 생성할 수 있는 기술을 포함하고 있다. 하지만 이는 수동으로 치조궁

(dental arch)의 경로를 그려주는 방법이 사용된다. 이와 같은 단점을 해결하기 위해, dental arch를 자동으로 검출하는 알고리즘에 대한 연구들이 진행되었다. T. Chanwimaluang [7]등은 치과용 아치 검출 시 자동화된 접근 방식을 제안하였지만, 처리하려는 영상에서 치아 사이의 빈 공간이 있는 경우 해당 알고리즘이 dental arch를 자동으로 검출할 수 없다는 단점이 있다. V. Sa-ing [8]등은 CT 영상에서 dental arch를 자동으로 감지하는 기능을 보완하여 앞선 문제들을 극복하였고, 해당 dental arch의 경로를 통해 CT에서 파노라마 영상 또한 자동으로 추출하는 알고리즘을 제안하였다. 이 연구를 통해 왜곡이 적은 시뮬레이션 파노라마 영상을 만들어 줄 수 있었지만, 실제 파노라마 X-선 영상과 같이 결과 영상에 치과 진단 시 불필요한 신체 부위들을 포함하고 있고, CT 영상이 없는 경우에는 해당 파노라마 영상을 생성해내지 못하는 큰 단점이 있었다.

위의 문제들을 해결하기 위해, 본 연구에서는 CT 영상에 의존하지 않고도 실제 파노라마 영상을 향상된 파노라마 영상으로 재 생성하는 심층 신경망 기반의 파노라마 영상 향상 시스템을 제안하였다. 해당 시스템 구축을 위해 생성기(generator) 네트워크 기반의 AD-ResUnet (Asymmetric Deep Residual U-net)을 제안하였고, 이 네트워크를 훈련시키기 위하여 파노라마 영상 데이터를 생성하는 파노라마 시뮬레이션 영상 생성 시스템(Panoramic image generator) 또한 설계하였다. 결과적으로 최종 훈련을 진행한 AD-ResUnet를 통해 실제의 파노라마 X-선 영상을 보다 진단에 용이한 파노라마 영상으로 향상시키는 영상 향상 시스템을 구축하였다.

II. 연구 방법

1. 데이터

데이터는 서울대 치과 병원에서 2019년 7월 3일부터 2020년 10월 17일까지 촬영된 3D Facial CT 영상을 사용하였다. 해당 연구는 서울대 치과 병원 의학 윤리 심의 위원회 (IRB:ERI20012)의 승인을 받았다. 해당 영상들은 0.6mm의 voxel로 구성되어 있고, Siemens (Germany) 사의 기기로 촬영되었다. 본 연구에서는 총 56명의 환자의 영상 중 무작위로 선택된 33명 환자들의 영상이 사용되었으며, 사용된 모든 데이터는 다중 검출기 전산화 단층 촬영 (Multi Detector Computed Tomography) 영상이다.

2. 파노라마 시물레이션 영상 생성 시스템

실제 파노라마 방사선 영상에서 치과 진단 시 불필요한 신체 부위들과 지방과 같은 연조직들을 자동으로 제거한 향상된 파노라마 X-선 영상으로 재생성하는 시스템을 구축하기 위해 심층 신경망을 적용하였다. 본 연구에서는 영상 향상 네트워크로 AD-ResUnet을 제안하였고, 해당 네트워크를 훈련시키기 위해서는 실제 파노라마 방사선 영상과 유사한 시물레이션 영상과, 치과 진단에 불필요한 부분들을 제거한 영상이 쌍 (paired) 형태로 요구된다. 이를 위해, CT를 사용하여 훈련에 필요한 두 종류의 영상 데이터를 쌍으로 획득하는 파노라마 영상 생성 시스템을 구축하였다.

파노라마 방사선 촬영 기기가 영상을 촬영할 때 사용하는 평균 광자

에너지 (Average Photon Energy; APE)는 CT 장비에서 사용하는 것과 다르다. 널리 사용되는 치과용 파노라마 방사선 기기에서는 30keV의 APE를 사용하고, 일반적인 CT 장비에서는 60keV의 APE를 사용하여 촬영한다[9]. 본 연구에서는 CT 영상을 파노라마 영상으로 변환하기 위해, 서로 다른 두 영상의 평균 광자 에너지에 의해 달라진 영상들의 픽셀 값(pixel value) 범위를 일치하도록 조정해주어야 한다. 이는 즉, 60keV의 APE에서 촬영된 CT 영상의 픽셀 값인 HU(Hounsfield Unit)을 30keV의 APE 일 때의 HU으로 변환해야 한다는 뜻이다. 이를 위해 식 2.2.1과 식 2.2.2를 통해, CT의 모든 HU 값을 선형 감쇠 계수 (Linear Attenuation Coefficient, μ)로 변환해야 한다.

$$HU_{60k} = 1000 \times \frac{\mu_{60k} - \mu_{water, 60k}}{\mu_{water, 60k} - \mu_{air, 60k}} \quad (2.2.1)$$

$$\mu_{60k} = \frac{HU_{60k} \times (\mu_{water, 60k} - \mu_{air, 60k})}{1000} + \mu_{water, 60k} \quad (2.2.2)$$

이때 식 2.2.2의 결과는 60keV APE에서의 선형 감쇠 계수가 되고, 선형 감쇠 계수는 질량 감쇠 계수에 밀도를 곱하여 얻을 수 있다[10]. CT를 식 2.2.1과 식 2.2.2를 통해 처리하여 얻는 결과인 60keV에서의 선형 감쇠 계수의 행렬을 식 2.2.3과 식 2.2.4를 통해 30keV의 선형 감쇠 계수들로 변환할 수 있다.

$$\mu_{30k, tissue} = 4.7723\mu_{60k}^4 - 14.45\mu_{60k}^3 + 1.5362\mu_{60k}^2 - 0.8329\mu_{60k} + 0.0007 \quad (2.2.3)$$

$$\mu_{30k, metal} = 0.0364\mu_{60k}^2 - 2.985\mu_{60k} + 3.3341 \quad (2.2.4)$$

선형 감쇠 계수를 변환하는 방정식은 60keV의 평균 광자 에너지에서의 선형 감쇠 계수와 30keV의 평균 광자 에너지에서의 선형 감쇠 계수 사이의 관계를 N차 방정식으로 근사하여 도출하였다. 해당 변환 방정식을 2가지로 나눈 것은 신체의 연조직 그룹과 경조직 및 금속 그룹간의 선형 감쇠 계수 값이 10배 이상 차이가 나기 때문에 하나의 그룹으로 포함시킬 경우, 변환 방정식의 오차가 증가하기 때문이다. 식 2.2.3 과 식 2.2.4 의 변환 방정식을 통해 30keV APE에서의 선형 감쇠 계수의 행렬로 변환한다.

재 생성된 CT 영상을 Z축을 따라 투영하고 수학적 형태학(Mathematical Morphology)의 한 종류인 얇게하기(Thinning)를 통해 Dental arch의 중간선을 찾는다. 중간선의 포물선을 따라 이동하면서 포물선 위의 점에 수직이 되는 선을 따라 CT 영상을 2차원 영상으로 샘플링(sampling)한다. 포물선 위의 점은 각각 x 값과 y 값의 변화량을 통해 4개의 구역으로 나누어 영상을 최대한 샘플링한다.

위의 방법으로 샘플링된 2차원 단면 영상들을 하나의 축으로 더하고, 그 결과인 1차원의 선들을 이어 붙여(concatenate) 파노라마 영상을 생성한다. 이때, 샘플링된 2차원 영상들에 처리해주는 방식에 따라 결과 영상의 종류가 정해진다. 설계된 파노라마 시뮬레이션 영상 생성 시스템의 결과 영상들을 실제의 파노라마 방사선 영상과 유사한 영상으로 만들어주기 위해 확대도 (Magnification) 개념을 적용하였다. 이는 샘플링된 2차원 단면 영상을 하나의 축으로 더해줄 때, 식 2.2.5를 따라 영상의 확대도를 적용하였다.

$$M = \frac{SOD}{SDD} \quad (2.2.5)$$

식 2.2.5의 SOD는 X-선원에서 물체까지의 거리(Source to Object Distance)이고, SDD는 X-선원에서 검출기까지의 거리(Source to Detector Distance)이고, M은 확대도(Magnification factor)를 의미한다. 앞서 서술한 바와 같이 2차원 단면 영상을 1차원 선으로 투영할 때 확대도를 적용할 경우, 그림 2.2.1 a~c 와 같이 실제 파노라마 X-선 영상과 유사하고 왜곡이 적은 시뮬레이션 파노라마 영상을 얻을 수 있었다. 또한 2차원 단면 영상들의 선형 감쇠 계수 값을 조절하여 특정 연조직을 제거하고, 치과 진단 시 불필요한 경추와 같은 불필요한 신체 부위들을 제거하여 관련 관심 영역만을 남기면, 그림 2.2.1 d~f 와 같은 향상된 파노라마 영상을 생성할 수 있다.

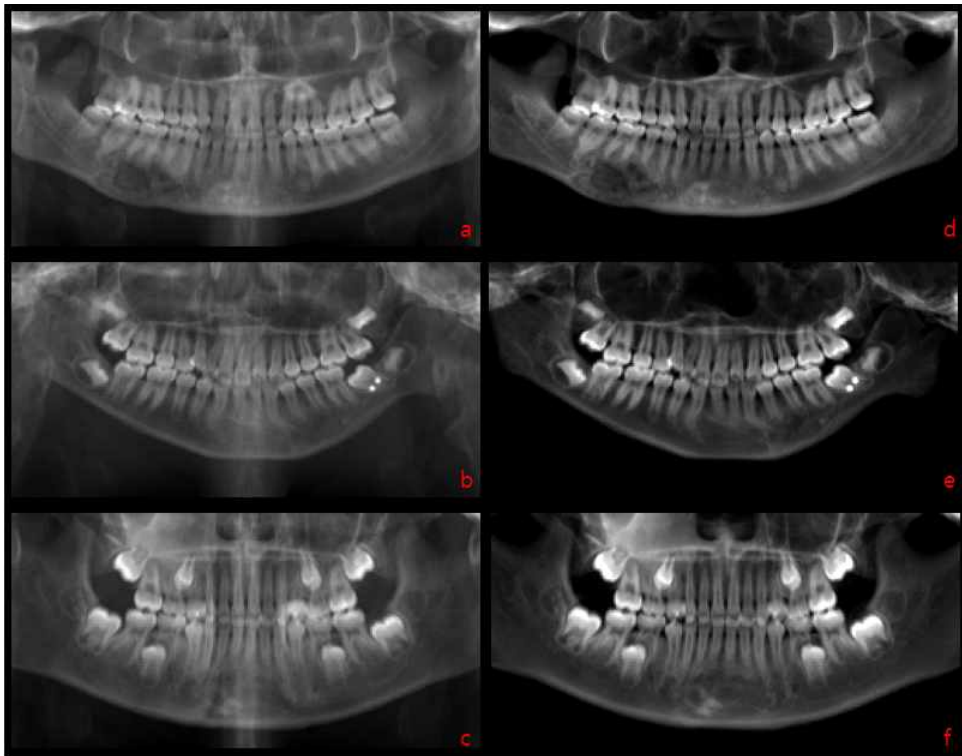


그림 2.2.1 panoramic image generator의 출력 영상

3. 제안하는 파노라마 영상 향상 네트워크

실제 파노라마 X-선 영상을 그림 2.2.1 d~f의 결과와 같이 치과 진단에 불필요한 연조직과 경추와 같은 경조직을 제거한 파노라마 영상으로 향상시키는 시스템을 효과적으로 구축하기 위해서 Deep-ResUnet[11]을 기반으로 하는 네트워크인 AD-ResUnet (Asymmetric Deep Residual Unet)을 제안하였다.

1) 관련 연구

(1) Residual Unit

다중 층 신경망(Multi-layer neural network)의 관점에서 심층 신경망이 깊어질수록 쉽게 성능이 향상될 수 있을 것 같지만, 오히려 훈련 오류를 증가시키거나 vanishing gradients 혹은 exploding gradients 문제가 발생될 가능성이 높아진다. 이런 성능 저하(degradation) 문제를 극복하기 위해 잔차 신경망(residual neural network) 구조를 제안하였고, 이는 동일한 깊이를 갖는 다른 네트워크들에 비해 좋은 성능을 보였고, 구조의 복잡성 또한 줄었다[12]. 여기에 사용한 residual neural network는 일련의 잔차 단위(residual unit) 들로 이루어져 있으며, 각 unit은 식 2.3.1 과 같이 표현이 된다.

$$\begin{aligned} y_k &= h(x_k) + F(x_k, W_k), \\ x_{k+1} &= f(y_k) \end{aligned} \quad (2.3.1)$$

x_k 와 x_{k+1} 이 각각 k 번째 잔차 단위의 입력과 출력이고, W_k 는 k 번째 층의 가중치일 때, 함수 $F(x_k, W_k)$ 는 잔차 함수(residual function),

$h(x_k)$ 는 입력으로 들어온 정보를 바로 출력 단으로 연결하는 identity mapping, $f(y_k)$ 는 ReLU 활성화 함수이다. 이 잔차 함수는 합성곱(convolution) 층과 활성화화(activation) 함수, 그리고 배치 정규화(Batch Normalization) 층의 다양한 조합에 따른 성능 차이를 갖는다. He 등이 진행한 연구[13]에서 제안한 pre-activation 구조가 post-activation 보다 네트워크 훈련에 있어 더 좋은 성능을 보였다. 이에 따라 그림 2.3.1과 같이 모든 layer의 각 잔차 단위들은 모두 pre-activation 구조를 갖도록 하였다.

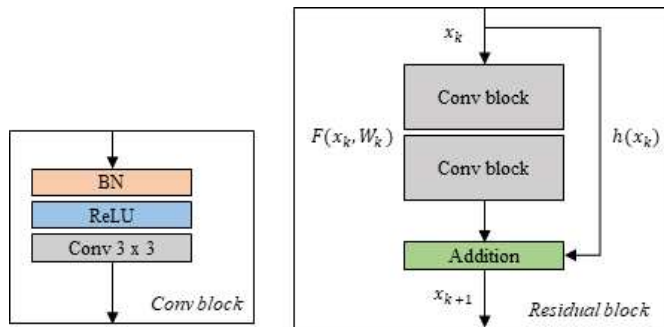


그림 2.3.1 pre-activation residual unit구조

(2) Deep ResUnet

Deep ResUnet의 구조는 U-Net[14]과 잔차 신경망을 결합함으로써 각각의 네트워크가 갖는 두 가지 장점들을 갖는 영상 분할을 위해 제안된 모델이다[11]. U-Net의 경우, 이미지 분할을 위해 제안된 완전 합성곱의 네트워크 (Fully Convolutional Network; FCN)을 기반으로 하는 모델로, 그림 2.3.2 과 같은 인코더-디코더 구조로 구성되어있다. Deep ResUnet

에서 사용된 U-Net의 스킵-연결(skip-connection) 구조를 통해 얇은 층(low-level)의 특징 맵을 깊은 층(high-level)의 특징 맵으로 성능 저하(degradation) 없이 구조적 정보를 직접 전파할 수 있다. 더욱이, 앞에서 설명한 바와 같이, Deep ResUnet은 모든 convolution 층마다 잔차 단위를 사용하여, 네트워크가 깊어지더라도 vanishing gradients 혹은 exploding gradients 문제없이 훈련 성능을 높였다.

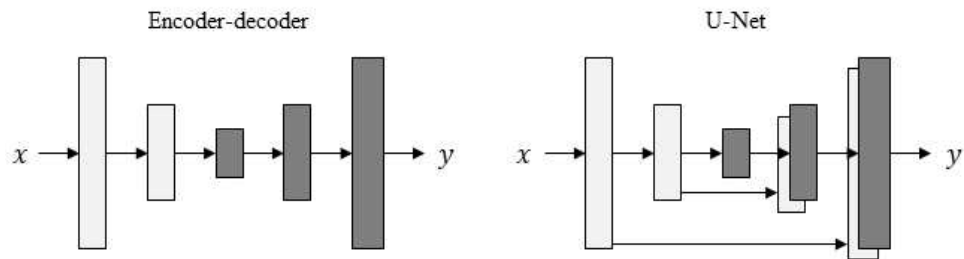


그림 2.3.2 Encoder-Decoder 구조와 U-Net 구조 비교

2) AD-ResUnet: 파노라마 영상 향상 네트워크

본 연구에서는 앞서 서술한 9-level의 Deep ResUnet을 기반으로 하는 AD-ResUnet을 제안하였다. 그림 2.3.3에서 볼 수 있듯이 제안한 네트워크는 Deep ResUnet의 encoder 앞단에 convolution layer들을 추가한 10-level 구조를 갖는다. 이는 배경 제거 층(background subtraction layer), 인코더(encoder), 다리(bridge), 디코더(decoder)를 포함하여 크게 4가지 부분으로 구성되어 있다.

Encoder에서 깊은 층의 특징 맵일수록 원본 데이터의 구조적 정보가 많이 잃는다. 이를 극복하기 위해, U-Net의 skip connection 구조(그림 2.3.2)를 추가하여, 해당 부분으로 low-level의 구조적 정보를 high-level

의 decoder로 손실이 없도록 전달 해주도록 하였다. Encoder의 skip connection을 통해 구조적 정보를 전달하기 이전에 최대한 많은 background 정보를 검출하여 제거할 수 있도록 첫 번째 층에 나타나 있는 것과 같이 합성곱 층(convolution layer)들을 추가 하였다.

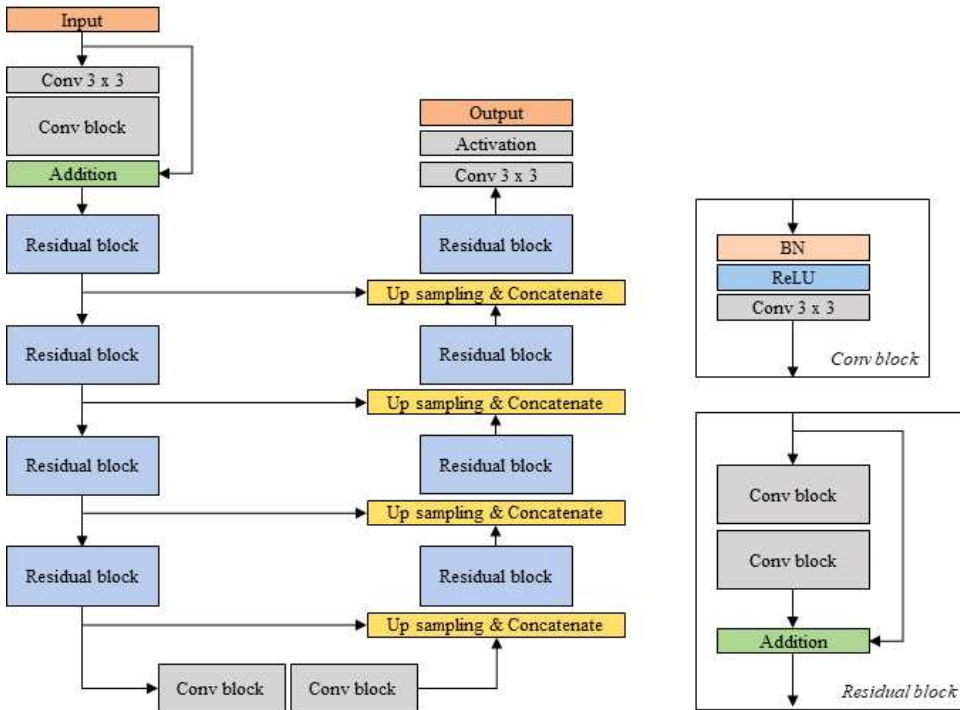


그림 2.3.3 본 연구에서 제안한 AD-ResNet 구조.

그림 2.3.3 에서 볼 수 있듯이, 잔차 블록(Residual block)의 합성곱 블록(Conv block)은 배치 정규화 층(Batch Normalization layer), ReLU 활성화 층(Activation layer), 그리고 3 x 3 합성곱 층(convolution layer)으로 이루어져 있다. 이때, background subtraction layer는 하나의 3 x 3

합성곱(convolution) 층과 하나의 Conv block, 그리고 해당 블록의 입출력에 직접 연결되는 identity mapping 으로 이루어져 있다.

Encoder는 4개의 잔차 블록(residual block)으로 구성되며, 각 블록은 pooling layer가 포함하지 않는다. 대신 각 잔차 블록의 첫 번째 convolution block의 stride를 2 로 지정하여 특징 맵을 절반으로 줄이는 방법을 사용하였다. 이는 pooling 작업을 pooling layer를 추가하지 않고도 fractionally strided-convolution을 통해 네트워크 모델을 훈련할 수 있고, 해당 방법은 pooling layer를 추가한 경우보다 모델의 표현력(expressiveness)을 향상시킬 수 있기 때문이다[15]. 또한 decoder는 4개의 잔차 블록으로 이루어져 있는데, 모든 블록의 앞단에는 이전 층의 특징 맵을 up-sampling 하였고, 해당 특징 맵을 encoder의 low-level에서 넘어온 특징 맵과 이어 붙였다. decoder의 최종 층에서는 Sigmoid 활성화 층과 1×1 convolution layer를 적용하여, 출력 영상의 크기를 입력 영상과 같도록 맞춰주었다. 제안한 네트워크의 모든 parameter들과 각 layer의 출력 크기는 표 2.1 와 같다. BN은 Batch Normalization, Conv는 Convolution, Residual은 Residual block을 의미한다. 또한 Add의 경우, 입력으로 들어오는 두 특징 맵을 더해주는 역할을 한다.

층		블록의 종류	레이어 변수	필터	보폭 크기	출력 크기
입력						128 x 256 x 1
Background subtraction layer	1	-	Conv		1	
		Conv block1	BN, ReLU, Conv	32 / 3 x 3	1	128 x 256 x 32
	2	Residual1	BN, ReLU, Conv	64 / 3 x 3	1	128 x 256 x 64
			Add (Conv block1, Residual1)	-		128 x 256 x 64
Encoder	3	Residual2	BN, ReLU, Conv	128 / 3 x 3	2	64 x 128 x 128
			BN, ReLU, Conv		1	
			Add (Residual1, Residual2)	-		64 x 128 x 128
	4	Residual3	BN, ReLU, Conv	256 / 3 x 3	2	32 x 64 x 256
			BN, ReLU, Conv		1	
			Add (Residual2, Residual3)	-		32 x 64 x 256
	5	Residual4	BN, ReLU, Conv	512 / 3 x 3	2	16 x 32 x 512
			BN, ReLU, Conv		1	
Bridge			Add (Residual3, Residual4)	-		16 x 32 x 512
	6	Conv block2	BN, ReLU, Conv	1024 / 3 x 3	2	8 x 16 x 1024
Decoder		Conv block3	BN, ReLU, Conv		1	
		Up-sampling1	-	1024 / 2 x 2		16 x 32 x 1024
		Concatenate1	Residual4, Up-sampling1			16 x 32 x 1536
	7	Residual5	BN, ReLU, Conv	512 / 3 x 3	1	16 x 32 x 512
			BN, ReLU, Conv		1	
			Add (Conv block3, Residual5)			16 x 32 x 512
		Up-sampling2	-	512 / 2 x 2		32 x 64 x 512
		Concatenate2	Residual3, Up-sampling2			32 x 64 x 768
	8	Residual6	BN, ReLU, Conv	256 / 3 x 3	1	32 x 64 x 256
			BN, ReLU, Conv		1	
			Add (Residual5, Residual6)			32 x 64 x 256
		Up-sampling3	-	256 / 2 x 2		64 x 128 x 256
		Concatenate3	Residual2, Up-sampling3			64 x 128 x 384
	9	Residual7	BN, ReLU, Conv	128 / 3 x 3	1	64 x 128 x 128
			BN, ReLU, Conv		1	
			Add (Residual6, Residual7)			64 x 128 x 128
		Up-sampling4	-	128 / 2 x 2		128 x 256 x 128
		Concatenate4	Residual1, Up-sampling4			128 x 256 x 192
	10	Residual8	BN, ReLU, Conv	64 / 3 x 3	1	128 x 256 x 64
			BN, ReLU, Conv		1	
			Add (Residual7, Residual8)			128 x 256 x 64
Output			Sigmoid, Conv	64 / 1 x 1	1	128 x 256 x 1

표 2.1 제안한 AD-ResUnet의 정보

4. 목적 함수

AD-ResUnet을 학습하기 위해 Mean Squared Error(MSE)와 Structural Similarity Index Map(SSIM), 2 종류의 목적 함수를 사용하였다. 전체 목적 함수는 식 2.4.1 과 같다.

$$L(y, G(x)) = \alpha \times MSE(y, G(x)) + \beta \times SSIM(y, G(x)) \quad (2.4.1)$$

MSE loss는 실수 기반의 결과에 대한 오차를 판별하는 목적함수이다. 제안한 파노라마 시뮬레이션 영상 생성 시스템의 결과 영상들을 각각 기본 파노라마 영상(x) 과 향상된 파노라마 영상(y) 이라고 할 때, 본 연구에서 제안한 영상 향상 네트워크인 AD-ResUnet을 통해 향상된 파노라마 영상을 $G(x)$ 라고 한다. 이때, MSE loss를 통해 $G(x)$ 와 y , $M \times N$ 크기의 두 영상 사이의 pixel 값의 차이를 식 2.4.2와 같이 표현한다.

$$L_{MSE}(y, G(x)) = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N (y_{ij} - G(x)_{ij})^2 \quad (2.4.2)$$

SSIM은 두 영상 사이의 유사성을 측정하기 위해 사용되는 잘 알려진 목적 함수이다. 인간의 시각 시스템(Human Visual System; HVS)을 토대로 개발되었으며, 기존의 오차 합계 방법이 아닌 휘도 왜곡, 대비 왜곡, 구조적 왜곡이라는 세 가지 요인들의 조합(식 2.4.3)으로 설계된다. 이를 통해, 본 연구에서 제안한 네트워크가 $G(x)$ 와 y 영상 간의 휘도, 대비, 구조에 대한 왜곡을 줄여주는 방향으로 훈련되도록 하였다. 이때, SSIM은 식 2.4.4 과 같이 표현된다.

$$\begin{cases} l(y, G(x)) = \frac{2\mu_y\mu_{G(x)} + c_1}{\mu_y^2 + \mu_{G(x)}^2 + c_1} \\ c(y, G(x)) = \frac{2\sigma_y\sigma_{G(x)} + c_2}{\sigma_y^2 + \sigma_{G(x)}^2 + c_2} \\ s(y, G(x)) = \frac{\sigma_y\mu_{G(x)} + c_3}{\sigma_y + \mu_{G(x)} + c_3} \end{cases} \quad (2.4.3)$$

$$\begin{aligned} SSIM(y, G(x)) &= l(y, G(x))c(y, G(x))s(y, G(x)) \\ &= \frac{2(\mu_y\mu_{G(x)} + c_1)(2\sigma_y\mu_{G(x)} + c_2)}{(\mu_y^2 + \mu_{G(x)}^2 + c_1)(\sigma_y^2 + \sigma_{G(x)}^2 + c_2)} \end{aligned} \quad (2.4.4)$$

제안한 네트워크인 AD-ResUnet을 훈련할 때 사용되는 식 2.4.1에서 각각의 MSE loss 와 SSIM 의 가중치인 α 와 β 를 하이퍼 파라미터 튜닝(Hyper-parameter tuning) 을 통해, 네트워크가 가장 적절한 방향으로 훈련될 수 있도록 정하였다.

5. 훈련 방법

본 연구에서는 앞선 전처리 과정에서 설계한 Panoramic Image Generator를 통해, 128 x 256 크기의 기본 시뮬레이션 파노라마 영상(x)과 이와 쌍(pair)인 향상된 파노라마 영상(y)을 생성해 내었다. 생성한 33개의 시뮬레이션 데이터 중 훈련용 데이터와 실험용 데이터로 각각 30개, 3개로 분리하였다. 훈련용 데이터 중 24개는 훈련에, 6개는 검증 데이터로 나누었다. 적은 훈련 데이터양을 보완하고자 기본 파노라마 영상(x)과 향상된 파노라마 영상(y)을 뺀 background($x-y$)의 대비를

조절하여 다시 향상된 파노라마 영상(y)와 더해주는 방식으로, 식 2.5.1 처럼 하나의 파노라마 영상 당 총 9개의 새로운 영상을 생성하였다.

$$x_k = ((x - y) \times k) + y, (1.1 \leq k \leq 1.9) \quad (2.5.1)$$

결과적으로 훈련 데이터는 총 216개가 사용되었고, 훈련 데이터의 다양성을 위해, random generator를 설계하여 좌우 반전, 회전, 평행이동, 비틀림 등의 효과를 주었다. 본 연구에서의 모든 과정은 표 2.2 과 같은 사양의 PC를 사용하여 진행하였다.

Method	Panoramic Image Generator	AD-ResUnet
CPU	Intel Core i7-9700	Intel Core i7-5930K
	(3.00GHz, 8 core, 8 threads)	(3.5 GHz, 6 cores, 12 threads)
GPU	NVIDIA GeForce RTX 2080	NVIDIA GeForce GTX TITAN X
	(8 GB)	(12 GB) \times 4
Memory	Samsung DDR4 2666	Samsung DDR4 2133
	(64GB = 16 x 4, 4/4)	(64 GB = 8 x 8)
Framework	Python 3.7.7	Tensorflow 2.1.0

표 2.2 사용한 하드웨어의 사양

6. 평가 지표

본 연구에서는 MSE(Mean Squared Error), PSNR(Peak Signal-to-Noise Ratio), SSIM(Structure Similarity)를 평가지표로 사용하였다. MSE는 비교할 두 영상에 대한 평균 제곱 오차를 뜻하며, 원본 영상(y)과 생성된 영상(\hat{y}) 사이의 픽셀 값들의 차이에 대한 측정값을 나타낸다. 비교할 두 영상의 크기가 $M \times N$ 일 때, 식 2.6.1과 같이 해당 함수를 나타낼 수 있다.

$$MSE(y, \hat{y}) = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N (y_{ij} - \hat{y}_{ij})^2 \quad (2.6.1)$$

PSNR의 경우, 식 2.6.2와 같이 표현되며, 영상이 가질 수 있는 최대 픽셀 값에 대한 잡음의 비를 나타내고, 생성된 영상(\hat{y})의 화질에 대한 손실 정도를 평가하기 위한 목적으로 사용하였다. 생성된 영상(\hat{y})이 복원되어야 할 원본 영상(y)과 비교 하였을 때 손실이 적을수록 높은 값을 갖으며, 0에 가까운 값일 경우 손실의 정도가 크다. 본 연구에서는 해당 지표의 입력 영상을 최대 픽셀 값으로 나누어, MAX 값을 1로 지정해 주었다.

$$PSNR(y, \hat{y}) = 10 \log_{10} \left(\frac{MAX^2}{MSE(y, \hat{y})} \right) \quad (2.6.2)$$

SSIM의 경우, 식 2.6.3 과 같이 표현되며, 원본 영상(y)과 생성된 영상(\hat{y}) 사이의 구조적 유사도와 인간의 관점에서의 시각적 화질 차이를 평가하기 위해 사용하였다.

$$SSIM(y, \hat{y}) = l(y, \hat{y})c(y, \hat{y})s(y, \hat{y}) \quad (2.6.3)$$

$$= \frac{2(\mu_y \mu_{\hat{y}} + c_1)(2\sigma_{y\hat{y}} + c_2)}{(\mu_y^2 + \mu_{\hat{y}}^2 + c_1)(\sigma_y^2 + \sigma_{\hat{y}}^2 + c_2)}$$

$$c_1 = (K_1 L)^2, c_2 = (K_2 L)^2, c_3 = \frac{c_2}{2}$$

식 2.6.3 에서 볼 수 있듯이, $l(y, \hat{y})$ 은 휘도(luminance), $c(y, \hat{y})$ 는 대비(contrast), 그리고 $s(y, \hat{y})$ 은 구조(structure)를 뜻한다. 해당 지표의 결과 값은 0 에서 1 사이의 값을 갖는데, 1 에 가까울수록 원본 영상(y)과 구조적 유사도를 갖는다. c_1 과 c_2 의 경우 0에 가까운 상수 값이며, 구성 인자인 L 은 입력 영상의 최대 픽셀 값을 의미하고, K_1 과 K_2 는 각각 0.01 과 0.03 이 사용되었다.

Ⅲ. 연구 결과

1. 하이퍼 파라미터 튜닝 결과

본 연구에서 제안한 AD-ResUnet의 훈련에 사용할 목적 함수인 MSE loss와 SSIM간의 가중치인 α , β 에 따른 훈련 결과를 비교하여, 하이퍼 파라미터(hyper-parameter)를 정하였다. 각 학습에 대하여 비교 지표로는 MSE loss, PSNR, 그리고 SSIM을 사용하였고, 가중치 변수인 $\alpha = (1, 5, 10, 20)$, $\beta = (1, 5, 10, 20)$ 에 대하여 총 8번의 훈련에 대한 결과를 비교하였다. 훈련 결과의 MSE가 낮을수록, PSNR과 SSIM은 높을수록 순위를 높게 책정하였고, 이들의 순위 합(Rank Sum) 이 가장 낮은 하이퍼 파라미터를 최종 훈련에 사용하였다. 하이퍼 파라미터 튜닝 결과는 표 4.1.1과 같고, 결과적으로 AD-ResUnet의 최종 훈련으로 $\alpha = 10$, $\beta = 1$ 을 사용하였다.

α	β	MSE	PSNR	SSIM	Rank Sum
1	1	0.00073	31.548	0.9455	14
5	1	0.00072	31.616	0.9459	9
10	1	0.00069	31.775	0.9466	3
20	1	0.00071	31.65	0.9459	7
1	5	0.00073	31.524	0.9458	14
1	10	0.00073	31.548	0.9462	10
1	20	0.00073	31.4	0.9452	18

표 3.1 Hyper-parameter tuning 결과 (AD-ResUnet)

2. 성능 비교

본 연구에서 제안한 AD-ResUnet의 성능을 평가하기 위해, 기존의 인코더-디코더(Encoder-Decoder)구조 생성기(generator)인 네트워크 중, U-Net 과 Deep ResUnet을 성능 비교에 사용하였다. 정확한 네트워크 성능을 비교하기 위해, U-Net과 Deep ResUnet은 AD-ResUnet의 Encoder-Decoder 경로와 동일하게 모두 9-level 구조로 설계하였고, 이들은 AD-ResUnet 의 최종 훈련과 동일한 훈련 데이터와 하이퍼 파라미터들을 통해 훈련을 진행하였다. 또한 모든 네트워크들은 검증 오차가 가장 작았던 epoch의 모델 가중치를 적용하였다. 이들의 훈련 결과를 비교하여 AD-ResUnet의 성능을 확인하였다.

최종 훈련된 U-Net, Deep-ResUnet, 그리고 AD-ResUnet을 통해, 훈련에 사용하지 않은 3개의 실험용 영상들(x)을 가상의 향상된 파노라마 영상(\hat{y})으로 재 생성하였다. 이 가상의 영상들을 실험용 데이터의 label인 향상된 시뮬레이션 파노라마 영상(y)과 비교하여 모든 평가 지표에 대한 결과를 추출하였다 (표 3.2).

모든 실험용 영상에 대한 전체적인 평가 지표의 경우, AD-ResUnet을 통해 생성된 가상의 향상된 파노라마 영상(\hat{y})이 U-Net 과 Deep-ResUnet을 통해 생성한 영상들에 비해 우수한 것을 볼 수 있었다(표 3.2). 조금 더 세부적인 영상 간 비교를 위하여 치과 진단 과정에서의 관심 영역(Region Of Interest; ROI)만을 앞선 방식과 동일한 방법으로 비교하였다. 그 결과 또한 AD-ResUnet의 모든 결과 영상들이 다른 네트워크들에 비해 우수하였다 (표 3.3).

		U-Net	Deep-ResUnet	AD-ResUnet
test image1	MSE	0.0006	0.00067	0.00052
	RSNR	32.2468	31.7317	32.8077
	SSIM	0.95058	0.95284	0.95554
test image2	MSE	0.00062	0.00065	0.00058
	RSNR	32.0666	31.8508	32.3486
	SSIM	0.95198	0.95003	0.95312
test image3	MSE	0.00117	0.00107	0.00096
	RSNR	29.3228	29.7245	30.1697
	SSIM	0.91951	0.9297	0.93107
average	MSE	0.0008	0.0008	0.00069
	RSNR	31.2121	31.1023	31.7753
	SSIM	0.94069	0.94419	0.94658

표 3.2 결과 영상들에 따른 네트워크별 성능 비교
(볼드체는 제안한 네트워크, 1순위는 파란색으로 표시함.)

		U-Net	Deep-ResUnet	AD-ResUnet
image1 ROI	MSE	0.0011	0.00129	0.00088
	RSNR	29.5732	28.8808	30.5519
	SSIM	0.94265	0.94395	0.95248
image2 ROI	MSE	0.0008	0.00078	0.0007
	RSNR	30.9778	31.0921	31.5739
	SSIM	0.9603	0.96297	0.96628
image3 ROI	MSE	0.00073	0.00061	0.00065
	RSNR	31.3955	32.177	31.8471
	SSIM	0.95048	0.95855	0.95605
average	MSE	0.00088	0.00089	0.00074
	RSNR	30.6489	30.7166	31.3243
	SSIM	0.95114	0.95516	0.95827

표 3.3 결과 영상들의 관심영역에 따른 네트워크별 성능 비교
(볼드체는 제안한 네트워크, 1순위는 파란색으로 표시함.)

그림 3.2.1, 그림 3.2.2, 그리고 그림 3.2.3 은 각 네트워크의 출력 영

상들(c ~ e)과 이들의 관심 영역에 대한 영상(f ~ i)을 시각화한 것이다. 관심 영역에 대한 영상간의 시각적 비교를 통해, 그림 3.2.1에서는 본 연구에서 제안한 AD-ResUnet의 출력 영상이 위 앞니(central incisor)의 뿌리 부분에 겹쳐진 경추(cervical vertebrae)의 세로선을 가장 잘 제거함을 볼 수 있었고, 그림 3.2.2에서는 AD-ResUnet의 출력 영상이 다른 네트워크의 출력 영상에 비해 위 앞니(central incisor) 부분을 구조적 손실 없이 가장 잘 복원함을 볼 수 있었다.

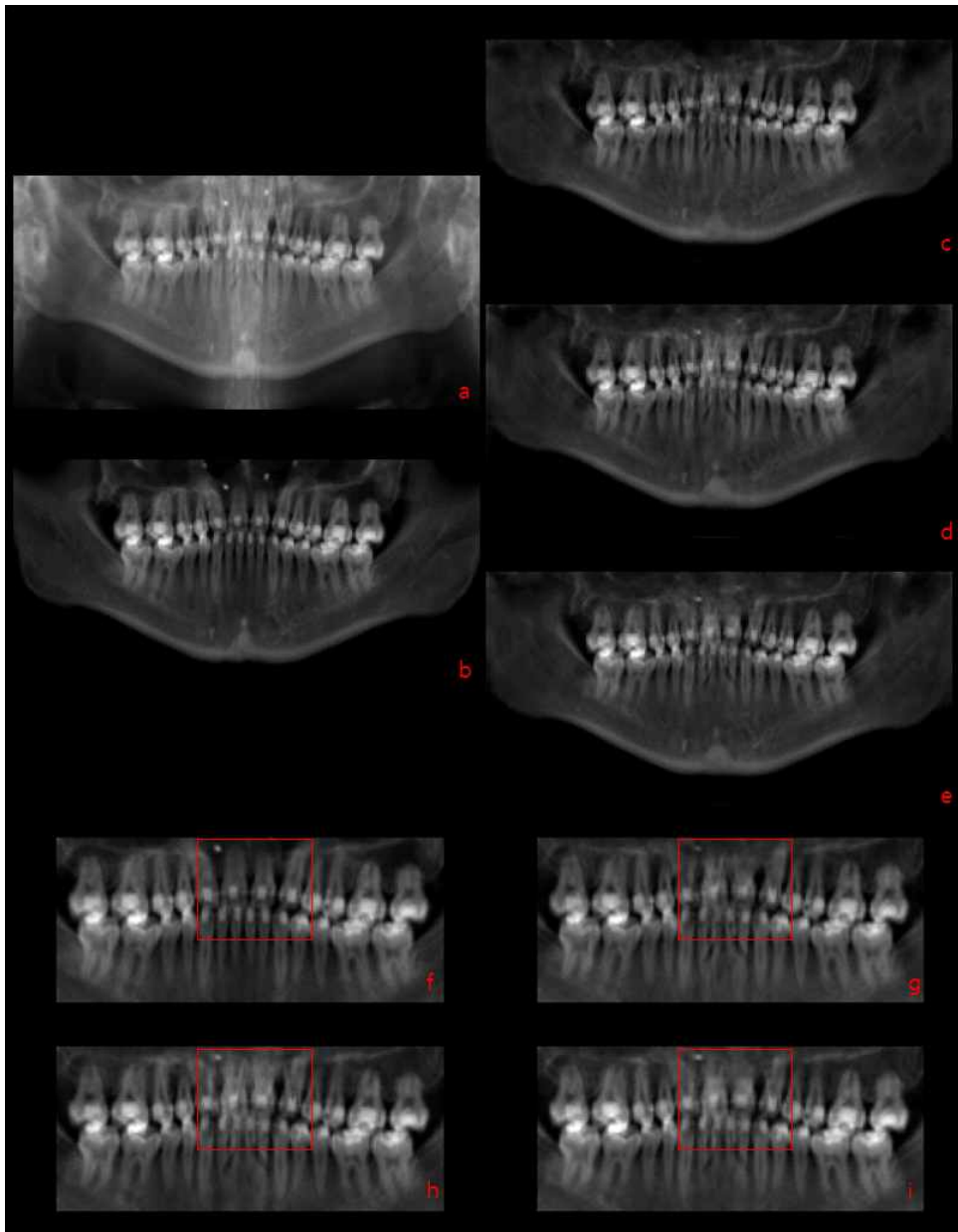


그림 3.2.1 a) test image1 b) label c) Unet의 출력 영상 d) Deep ResUnet의 출력 영상 e) AD-ResUnet의 출력 영상 f) b의 관심영역 g) c의 관심영역 h) d의 관심영역 i) e의 관심영역

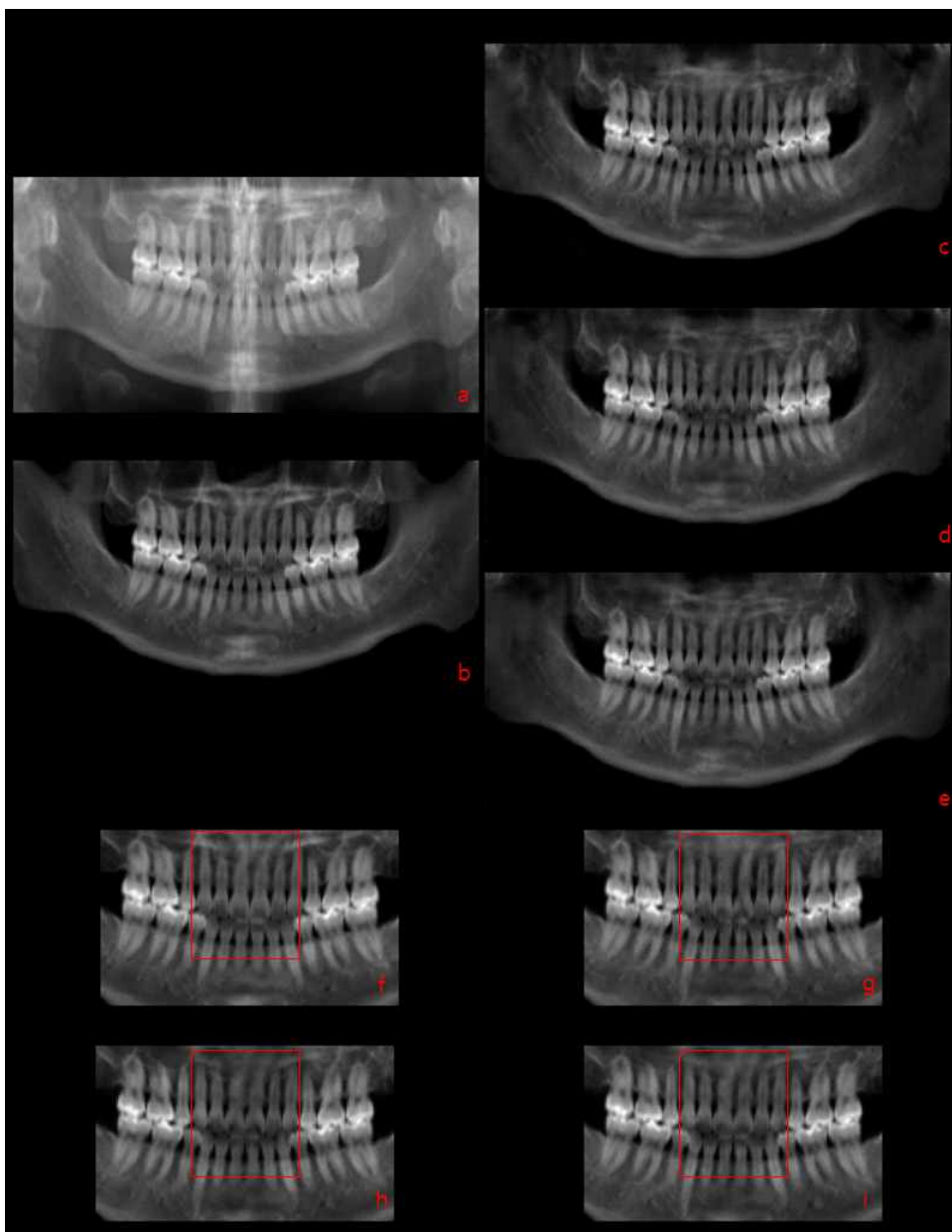


그림 3.2.2 a) test image2 b) label c) Unet의 출력 영상 d) Deep ResUnet의 출력 영상 e) AD-ResUnet의 출력 영상 f) b의 관심영역 g) c의 관심영역 h) d의 관심영역 i) e의 관심영역

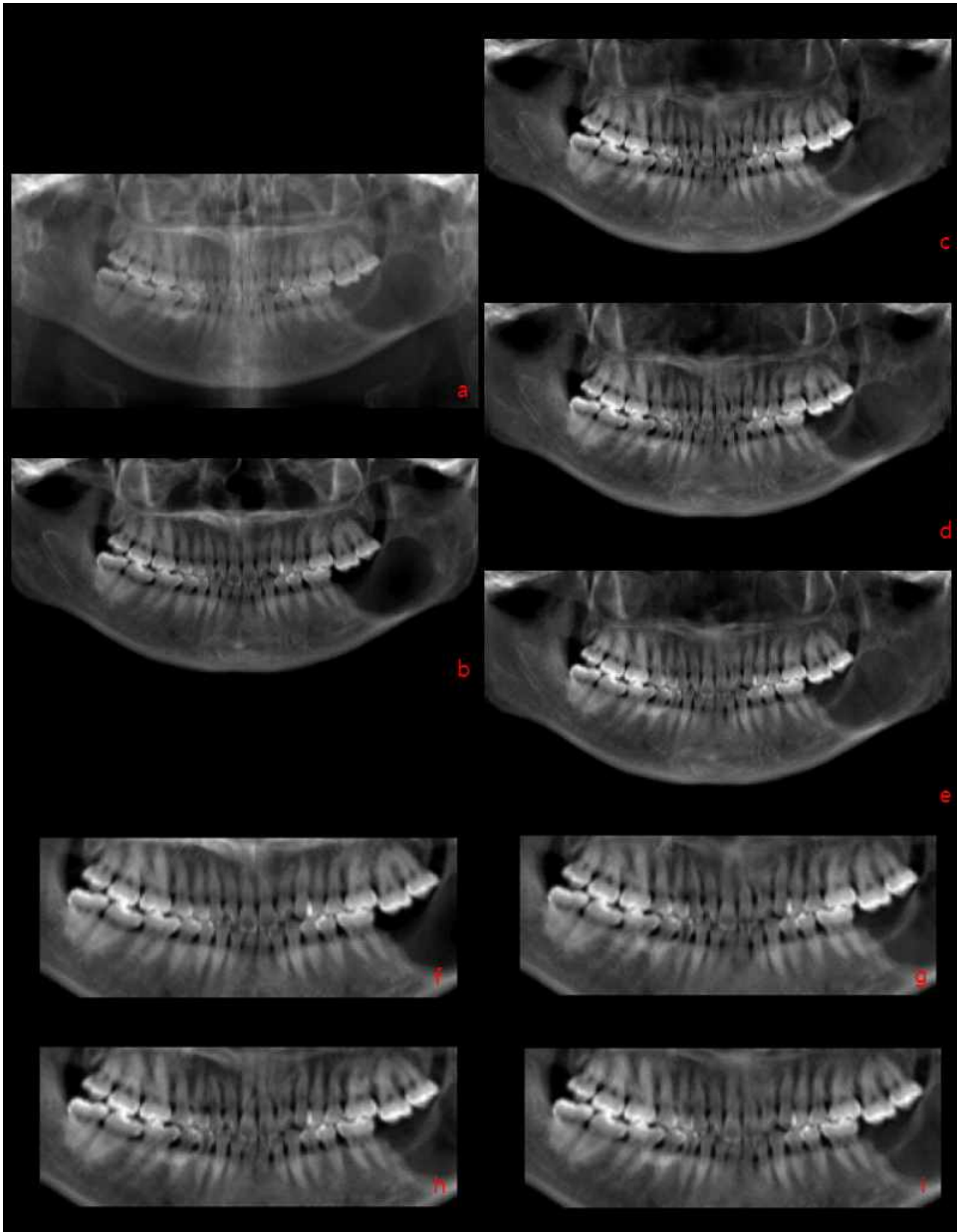


그림 3.2.3 a) test image3 b) label c) Unet의 출력 영상 d) Deep ResUnet의 출력 영상 e) AD-ResUnet의 출력 영상 f) b의 관심영역 g) c의 관심영역 h) d의 관심영역 i) e의 관심영역

3. 실제 파노라마 영상에 대한 네트워크의 성능

훈련에 사용되지 않은 실험용 데이터를 통해 본 연구에서 제안한 AD-ResUnet의 영상 향상 성능이 다른 네트워크들에 비해 우수한 것을 확인 할 수 있었다. 하지만, 이 연구의 최종 목적은 실제 치과 파노라마 X-선 영상을 개선하는 시스템을 개발하는 것이다. 즉, 실제 파노라마 X-선 영상이 최종 훈련된 AD-ResUnet에 입력될 때 출력되는 영상을 통해 성능을 평가해야 한다는 의미를 갖는다. 이를 위해, 네트워크 입력에 맞게 실제 파노라마 영상의 의미 없는 정보를 잘라낸 뒤, 크기를 조정해 다음 AD-ResUnet을 통해 처리하였다. 처리된 실제 파노라마 X-선 영상, 총 26개를 얻었으며, 이 중 무작위로 5개의 영상을 시각화하였다 (그림 3.3.1).

그 결과, 그림 3.3.1 에 나타나 있듯이 AD-ResUnet에 입력했던 실제 파노라마 X-선 영상에 대해서, 겹쳐진 경추(Cervical Vertebrae)를 잘 제거하였고, 영상 전체를 blur하게 만들던 연조직으로 인한 noise를 잘 제거하는 것을 확인 할 수 있었다.

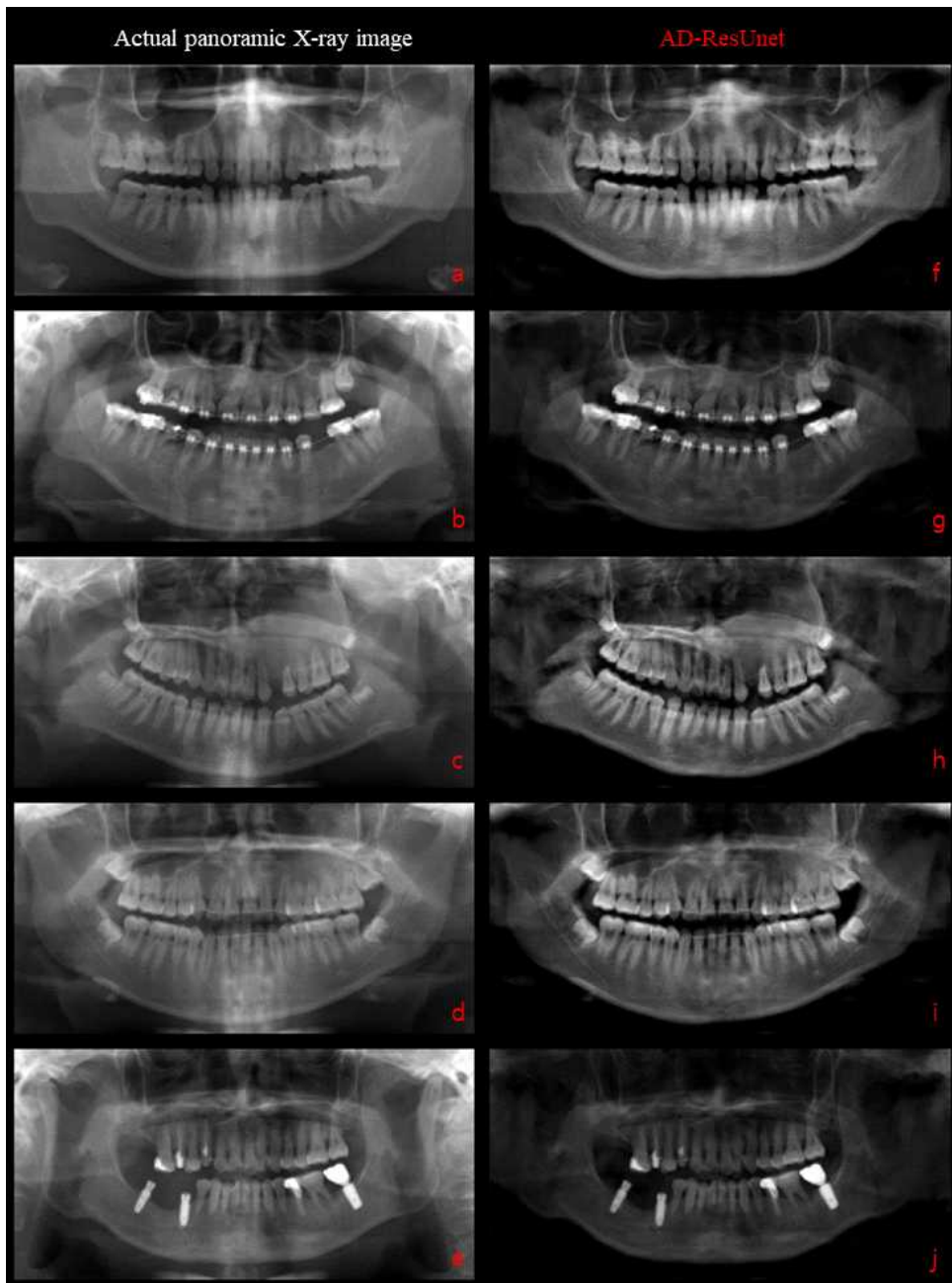


그림 3.3.1 AD-ResUnet을 통해 처리한 실험용 영상과 결과 영상

IV. 결론

본 연구는 CT 영상을 사용하지 않고도 기존의 치과 파노라마 X-선 영상을 진단 시 도움이 되도록 개선된 파노라마 영상으로 재 생성하는 영상 향상 시스템을 구축하는 것을 목표로 하였으며, 이를 위해 심층 신경망 기반의 AD-ResUnet 이라는 영상 향상 네트워크를 제안하였다. 네트워크를 훈련시키기 위해서는 실제 치과 파노라마 X-선 영상과 해당 영상과 쌍(pair)을 이루는 개선된 파노라마 X-선 영상을 필요로 한다. 이러한 쌍 데이터는 기존의 치과용 파노라마 X-선 장비를 통해 획득 할 수 없으며, 실제 X-선 영상에서 치아와 겹쳐진, 불필요한 부분들을 직접 제거하는 것 또한 어려운 일이다.

이 문제들을 해결하기 위해, CT 영상을 사용하여 실제 치과 파노라마 X-선 영상과 유사한 2D 파노라마 영상을 생성하는 동시에, 치과 진료에 불필요한 신체 부위들이 제거된 개선된 파노라마 영상을 생성하는 파노라마 영상 생성 시스템(panoramic image generator)을 구축하였다. 해당 시스템을 통해 획득한 두 가지 유형의 영상들을 사용하여, CT에 의존하지 않고도 실제 치과 파노라마 영상을 개선된 영상으로 향상시키도록 AD-ResUnet을 훈련하였다. 이 네트워크는 동일한 조건 하에 훈련한 다른 네트워크들에 비해 영상 향상 능력이 우수한 것을 확인 할 수 있었다.

그러나 이 연구의 최종 목적은 실제 치과 파노라마 X-선 영상을 향상시키는 시스템을 개발하는 것이다. 그러므로 최종 훈련된 AD-ResUnet이 파노라마 영상 생성 시스템(panoramic image generator)을 통해 획득한 시뮬레이션 영상뿐만 아니라, 실제 파노라마 영상에서도

영상을 향상시키는 기능이 제대로 작동하는 지 그 여부를 확인하였다.

이를 위해, 총 26개의 실제 파노라마 영상에 대하여 AD-ResUnet을 통해 결과를 획득하였으며, 실제 영상에서도 중심부에 치아와 겹쳐진 경추(Cervical Vertebrae)를 잘 제거하였고, 영상을 흐릿하게 하는 연조 직들 또한 잘 제거한 것을 확인할 수 있었다. 하지만 그림 4.3.1에서 볼 수 있듯이, 몇몇의 영상에서는 겹쳐진 뼈들을 완벽하게 제거해내지 못하여 하악골 양쪽에 겹쳐짐 현상이 남아 있는 것을 확인 할 수 있었다. 이는 네트워크 훈련을 위해 사용한 데이터의 다양성과 양이 부족했기 때문에 발생하는 훈련 데이터의 일반화(generalization) 문제이며, 또한 실제 파노라마 영상이 AD-ResUnet의 훈련에 직접적으로 포함되어있지 않았기 때문에 발생한 것으로 사료된다.

이러한 문제들을 해결하기 위해 향후의 연구는 파노라마 영상 생성 시스템(panoramic image generator)이 실제 파노라마와 더 유사한 영상을 생성할 수 있도록 시스템 설계 시 적용했던 매개변수들을 보완할 것이다. 또한, 더 다양한 데이터를 생성할 수 있도록 기본 파노라마 영상(x_n)과 향상된 파노라마 영상(y_n)을 뺀 $\text{background}(x_n - y_n)$ 의 대비를 조절하여 다시 다른 향상된 파노라마 영상들에($y_{1,2,\dots,k}$)와 더해주는 방식을 사용할 것이다.

본 연구에서 제안한 네트워크인, AD-ResUnet은 하나의 생성기(generator) 네트워크만으로 이루어져 있다. 후속 연구에서는 생성기(generator) 네트워크와 판별기(discriminator) 네트워크가 서로 경쟁하며 훈련 성능을 향상시키는 생산적 적대 신경망(Generative Adversarial Network; GAN)[16]을 적용함으로써 영상 향상 시스템의 성능을 향상시킬 것이다.

V.참고 문헌

- [1] P. Rajpurkar et al., “Deep learning for chest radiograph diagnosis: A retrospective comparison of the CheXNeXt algorithm to practicing radiologists,” PLoS Med, vol. 15, no. 11, p. e1002686, Nov. 2018, doi: 10.1371/journal.pmed.1002686.
- [2] J. Krois et al., “Deep Learning for the Radiographic Detection of Periodontal Bone Loss,” Scientific Reports, vol. 9, no. 1, Art. no. 1, Jun. 2019, doi: 10.1038/s41598-019-44839-3.
- [3] R. A. Danforth and D. E. Clark, “Effective dose from radiation absorbed during a panoramic examination with a new generation machine,” Oral Surg Oral Med Oral Pathol Oral Radiol Endod, vol. 89, no. 2, pp. 236-243, Feb. 2000, doi: 10.1067/moe.2000.103526.
- [4] F. Ezoddini Ardakani, M. Zangouie Booshehri, and B. Behniafar, “Evaluation of the Distortion Rate of Panoramic and Periapical Radiographs in Erupted Third Molar Inclination,” Iran J Radiol, vol. 8, no. 1, pp. 15-21, Mar. 2011.
- [5] Y. Hayakawa, M. Wakoh, H. Fujimori, Y. Ohta, and K. Kuroyanagi, “Morphometric analysis of image distortion with rotational panoramic radiography,” Bull Tokyo Dent Coll, vol. 34, no. 2, pp. 51-58, May 1993.
- [6] R. A. Kayal, “Distortion of digital panoramic radiographs used for implant site assessment,” J Orthod Sci, vol. 5, no. 4, pp. 117-120,

- 2016, doi: 10.4103/2278-0203.192113.
- [7] T. Chanwimaluang, S. Sotthivirat, and W. Sinthupinyo, “Automated dental arch detection using computed tomography images,” in 2008 9th International Conference on Signal Processing, Oct. 2008, pp. 737–740, doi: 10.1109/ICOSP.2008.4697235.
- [8] V. Sa-ing, K. Wangkaoom, and S. S. Thongvigitmanee, “Automatic dental arch detection and panoramic image synthesis from CT images,” in 2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Jul. 2013, pp. 6099–6102, doi: 10.1109/EMBC.2013.6610944.
- [9] “X-rays and ionizing radiation,” Design World, Dec. 02, 2015. <http://www.testandmeasurementtips.com/new-articles/8554> (accessed Dec.21, 2020).
- [10] curtis.suplee@nist.gov, “X-Ray Mass Attenuation Coefficients,” NIST, Sep. 17, 2009. <https://www.nist.gov/pml/x-ray-mass-attenuation-coefficients> (accessed Dec. 21, 2020).
- [11] Z. Zhang, Q. Liu, and Y. Wang, “Road Extraction by Deep Residual U-Net,” IEEE Geoscience and Remote Sensing Letters, vol. 15, no. 5, pp. 749–753, May 2018, doi: 10.1109/LGRS.2018.2802944.
- [12] K. He, X. Zhang, S. Ren, and J. Sun, “Deep Residual Learning for Image Recognition,” in 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Jun. 2016, pp. 770–778, doi:

10.1109/CVPR.2016.90.

- [13] K. He, X. Zhang, S. Ren, and J. Sun, “Identity Mappings in Deep Residual Networks,” in *Computer Vision – ECCV 2016*, Cham, 2016, pp. 630–645, doi: 10.1007/978-3-319-46493-0_38.
- [14] O. Ronneberger, P. Fischer, and T. Brox, “U-Net: Convolutional Networks for Biomedical Image Segmentation,” in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, Cham, 2015, pp. 234–241, doi: 10.1007/978-3-319-24574-4_28.
- [15] J. T. Springenberg, A. Dosovitskiy, T. Brox, and M. Riedmiller, “Striving for Simplicity: The All Convolutional Net,” arXiv:1412.6806 [cs], Apr. 2015, Accessed: Dec. 21, 2020. [Online]. Available: <http://arxiv.org/abs/1412.6806>.
- [16] I. J. Goodfellow et al., “Generative Adversarial Networks,” arXiv:1406.2661 [cs, stat], Jun. 2014, Accessed: Dec. 21, 2020. [Online]. Available: <http://arxiv.org/abs/1406.2661>.

Abstract

AD-ResUNet : Development of Panoramic Dental X-ray Image Improving System using Deep Learning Network

Kang, Hee Guen

Advised by Prof. Son, Young Don

Dept. of Health Sciences and Technology, GAIHST,
Gachon University, Incheon, Republic of Korea

Panoramic radiographs are the most basic and important tool for most dental practices. Panoramic dental X-ray images, however, have the disadvantage of distortion compare with the dentition of humans and include unnecessary body parts that interfere with a dental practice, such as cervical vertebrae and soft tissue. To solve this problem, many studies have been conducted related to the construction of automated systems that generate low distortion simulated panoramic images via computed tomography images. However, this in turn does not improve actual panoramic images but relies only on CT to generate new simulation images. In this study, I proposed AD-ResUnet that regenerates actual panoramic images into

improved panoramic images to address these problems. This network is built based on encoder-decoder generator networks, and all layers of the network applied a pre-activation residual block. The network was trained by two loss functions: MSE and SSIM. I evaluate the performance of AD-ResUnet by comparing the results of test data processed with other generator models trained under the same conditions. Comparison metrics include MSE, PSNR, SSIM, which resulted in AD-ResUnet outperformed comparable networks at 0.00069, 31.7753, and 0.94658, respectively.

감사의 글

저는 처음 연구실을 하드웨어 팀으로서 학부생 때부터 석사과정 1년 차까지 2년 반을 보냈었습니다. 그러다 인공지능과 의료영상이라는 것에 호기심을 느끼고 접한 것은 석사과정 2년차부터였고 남들보다 느렸던 시작으로 같은 연구실 사람들에 비해 과제를 해나가기엔 많이 부족하고 모르는 것투성이였습니다. 힘든 날들이 많았지만, 그래도 연구를 잘 마무리 할 수 있도록 도와준 모든 제 주변사람들에게 감사의 말을 전하고 싶습니다.

학부생 때부터 지금까지 거의 4년을 연구실에서 지내며 선생님처럼 많은 것들을 가르쳐주셨고, 아플 때는 가족처럼 그 누구보다 걱정해주셨고, 연구에 막힘이 있을 때 언제나 도와주시고 믿어주시던 제 지도교수님인 손영돈 교수님께 가장 먼저 감사의 인사를 드리고 싶습니다. 교수님의 지원과 조언, 그리고 믿음이 있었기에 방황하던 시기를 잘 넘기고 인공지능 분야로 석사 과정을 마칠 수 있었습니다. 치과영상 분야에 관심을 갖게 해주셨고, 실무적인 조언들을 많이 해주셨던 김행근 교수님께도 감사드립니다. 학위 논문 심사에 응해주시고 제 연구에 대해 좋은 말씀을 아낌없이 해주셨던 조장희 교수님께도 감사드립니다.

우리 동기들, 태진이형, 기범이형에게도 감사의 말을 전합니다. 동기라는 이유로 장난을 많이 치고 짓궂게 대했지만, 태진이형은 인공지능과 의료 영상 분야에 있어 저에게 스승으로서, 기범이형은 하드웨어 팀에서 오래 같이 일한 친구로서 정말 감사하다는 말을 전하고 싶었습니다.

MMMIL 연구실 학생들에게도 감사의 말을 전합니다. 용기형에게 항상

건강 잘 챙기고, 말은 연구 잘 진행하여 좋은 결실을 맺었으면 좋겠습니다. 이제 입학하는 지현이에게 제 연구 주제를 이어나가는 주는 것에 대해 고맙다는 말과 함께 앞으로의 생활에 응원을 전하고 싶습니다. 그리고 연구 분야가 겹치지 않아 많은 시간 함께하지 못한 김정희 교수님과, 학부생 연구원들, 송환, 예린, 한봄, 상민, 한나, 민지에게도 응원과 고맙다는 말을 전하고 싶습니다. MMMIL 연구실의 모든 학생들에게 앞으로의 연구실 생활이 각자에게 소중하고 의미 있는 시간이 되었으면 좋겠습니다. 남는 것은 지식과 사람인 것 같습니다. 꼭 두 마리 토끼를 다 잡으셨으면 좋겠습니다.

그리고 선배이자 동생인 봉혁이에게도 감사의 말을 전합니다. 고맙다는 말 한번으로는 표현하지 못할 만큼 진심으로 항상 고마운 마음을 품고 있습니다. 동생이지만 항상 연구실의 선배로서 무거운 책임감을 혼자 짊어지려는 봉혁이를 많이 도와주지 못해 미안했다는 말도 전하고 싶습니다. 또한 오랜 시간 함께해왔던 동료로서 봉혁, 보경, 재열이형, 재은이에게 고맙다는 말과, 항상 건강 챙기자는 말을 전하고 싶습니다.

마지막으로 항상 응원해주신 우리 가족, 아버지, 어머니, 누나, 할아버지, 할머니, 그리고 외할머니께 감사드립니다. 아버지, 어머니가 저를 항상 든든하게 지탱해주셨기에 제가 힘이 났습니다. 우리 가족 모두 사랑합니다.

많은 사랑과 응원, 지원이 있었기에 제가 연구실에서 의미 있는 시간을 보내며, 석사 과정을 무사히 마칠 수 있었습니다. 제가 글로 남기지 않았지만 저를 응원해주셨던 모두에게 감사의 말을 전하고 싶습니다. 감사합니다.

강 희 근 올림

AD-ResUnet: 심층 신경망을 이용한
치과 파노라마 X-선 영상 향상 기술 개발

姜 熙 根