

차세대 디지털 병리를 위한 Label Free 디지털염색 알고리즘 비교 연구

황석민¹, 정연우¹, 김동범¹, 이승아², 조남훈³, 이종하^{*}

¹계명대학교 공과대학 의용공학과

²연세대학교 공과대학 전기전자공학부

³연세대학교 의과대학 병리학교실

The Novel Label Free Staining Algorithm in Digital Pathology

Seok-Min Hwang¹, Yeun-Woo Jung¹, Dong-Bum Kim¹, Seung Ah Lee²,

Nam Hoon Cho³, and Jong-Ha Lee^{1*}

¹Dept. of Biomedical Engineering, School of Medicine, Keimyung University

²Dept. of Electrical and Electronic Engineering, Yonsei University

³Dept. of Pathology, College of Medicine, Yonsei University

요약 암세포와 정상세포를 구분하기 위해서는 H&E(Hematoxylin&Eosin) 염색이 필요하다. 병리 염색은 많은 비용과 시간이 필요하다. 최근 이러한 비용과 시간을 줄이고자 디지털 염색 방법이 소개되고 있다. 본 연구에서는 병리 H&E 염색의 디지털 변환 방법에 대한 새로운 알고리즘을 제안한다. 첫 번째 알고리즘은 Pair방법이다. 본 방법은 FPM(Fourier Ptychographic Microscopy)으로 촬영된 염색된 Phase 영상과 염색되지 않은 Amplitude 영상을 학습하여 염색된 Amplitude 영상으로 변환한다. 두 번째 알고리즘은 Unpair방법이다. 본 방법은 염색된 형광현미경 영상과 염색되지 않은 형광현미경 영상을 학습하여 모델링하여 디지털 염색을 수행한다. 본 연구에서는 GAN(generative Adversarial Network)를 활용하여 디지털 염색을 진행하였다. 연구 결과 Pair방법과 Unpair방법 모두 우수한 성능의 디지털 염색 결과를 확보하였다.

• **주제어** : 인공지능, 디지털병리, 컴퓨터보조진단, 데이터처리, 디지털염색

Abstract To distinguish cancer cells from normal cells, H&E (Hematoxylin & Eosin) staining is required. Pathological staining requires a lot of money and time. Recently, a digital dyeing method has been introduced to reduce such cost and time. In this paper, we propose a novel digital pathology algorithms. The first algorithm is the Pair method. This method learns the dyed phase image and unstained amplitude image taken by FPM (Fourier Ptychographic Microscopy) and converts it into a dyed amplitude image. The second algorithm is the unpair method. This method use the stained and unstained fluorescence microscopic images for modeling. In this study, digital staining was performed using a generative adversarial network (GAN). From the experimental results, we noticed that both the pair and unpair algorithms shows the excellent performance.

• **Key Words** : Artificial Intelligence, Digital Pathology, Computer Aided Diagnostic, Data Processing

Received 06 February 2023, Revised 27 March 2023, Accepted 30 March 2023

* **Corresponding Author** Jong-Ha Lee, Dept. of Biomedical Engineering, 1079 Dalgobeoldaero, Keimyung Univeristy, South Korea.
E-mail: segeberg@kmu.ac.kr

I. 서론

인공지능은 빅데이터를 이용해 학습하고 고유의 패턴을 인식하는 기술이다. 이전의 영상 필터를 통한 단순 변환이 아닌 딥러닝 기반 인공지능 모델링 기법도 도입됨으로써 많은 분야에 있어 새로운 패러다임을 형성했다. 그중에서도 유전 복잡성이 DNA 복사에 주는 영향을 예측하는 부분이나 의료 분야에 있어 암이나 종양 같은 부분을 영상에서 찾는 분야에서도 높은 성능을 보여주는 논문들이 소개되고 있다 [1]. 최근 병리학 분야에서도 인공지능 기술 등을 활용한 접근이 많이 소개되고 있다. 디지털 염색이란 세포 염색이 되지 않은 슬라이드를 디지털로 염색하는 방법이며 많은 알고리즘이 소개되고 있다[2]. 판독 슬라이드에 가장 많이 사용되는 세포염색 방법은 헤마톡실린과 에오신 염색 (Hematoxylin staining & Eosin staining, H&E) 방법이다. 병리학 세포 판독 염색에 있어 약 80%가 H&E 염색법을 이용하고 있다는 보고가 있다[3]. 두 염색 방법은 세포의 다른 부분을 염색하는데 헤마톡실린 염색은 세포핵을 염색하고 에오신은 세포질을 염색한다[4]. 또한 최근에는 위상차 현미경이 병리학 분야에서 소개되고 있으며 기존 형광 현미경에서 볼 수 없는 다양한 특징점들을 위상차 현미경에서 추출해 진단에 활용하고 있다.

위상차 현미경이란 아주 작은 빛을 산란하고 흡수하는 샘플을 찍는 기법을 활용하는 현미경이다. 카메라와 같은 기기로 이미지를 얻을 수 있으며 아주 작은 빛의 산란과 흡수를 이용해 각 샘플의 고유한 영상 대비를 생성할 수 있다. 생성된 영상 대비는 대상체의 특정한 진폭과 위상을 나타내며 이러한 진폭과 위상을 변화시켜 대상체의 아주 얇은 두께와 굴절률을 파악할 수도 있는 논문도 소개되고 있다[5-6]. 위상차 현미경으로 이미지를 획득하면 시료를 염색할 필요가 없으므로 오랜 시간 동안 살아있는 세포의 관찰이 가능하여 세포의 행동 패턴 분석에 적합하다는 논문도 제시되어 있다[7]. 정량적 위상 이미징의 돋보이는 특징 중 하나는 극도로 민감한 빛의 산란 데이터를 생성할 수 있다는 것이다. 이는 주어진 평면에서 복잡한 필드에 대한 지식을 통해 원거리 영역을 포함한 다른 평면에서의 필드 분포를 유추할 수 있기 때문이다. 푸리에 변환 광 산란이라고 불리는 이 접근법은 일반적인 각도계

기반 각도 산란보다 훨씬 민감하다. 결과적으로 푸리에 변환 광 산란은 미세한 세포 하위 구조의 산란 특성을 쉽게 렌더링할 수 있다고 소개되고 있다[8]. 일반적인 광학현미경으로 촬영 시에는 위상 차이를 명암 차이로 변화시켜 이미지가 나오는 것으로 알려져 있다 [9-10]. 그러나 명암 차이로 이미지를 보는 광학현미경은 세포의 특징 등을 보기 위한 H&E 염색이 필수적이다[11]. 본 연구에서는 디지털 염색을 위해 위상차 현미경의 일종인 푸리에 타이코그래픽 현미경(Fourier Ptychographic Microscopy, FPM)을 이용하고자 한다. 푸리에 타이코그래픽 현미경은 저 해상도의 푸리에 공간의 샘플 강도 영상을 통해 매개변수로 Phase image와 Amplitude image를 얻을 수 있다. Phase image는 시료를 염색하지 않아도 세포를 빛의 산란 데이터로 생성이 가능하고 분석하기 적합해 H&E 염색이 되지 않은 세포도 다양한 형태의 특징 등을 예측할 수 있는 논문들이 소개되고 있다[12-13].

본 논문에서는 H&E 염색이 되지 않은 세포를 H&E 디지털 염색을 하는 것을 목표로 한다. 그 방법으로 두 가지 방법을 제시하는데 먼저 Pair 방법의 경우에는 FPM으로 촬영한 슬라이드에서 phase image를 통해 amplitude image로 변환하는 디지털 염색 알고리즘을 적용한다. Unpair방법의 경우에는 형광 현미경으로 촬영한 염색이 되지 않은 image와 염색이 된 image를 모델링하여 디지털 염색을 하는 방법 등을 제안한다. 본 제안 방법은 H&E 염색의 염색 증발과 시간, 비용을 절약할 수 있을 것으로 기대한다.

II. 연구과정

2.1 Pair Method

Pair 방법은 염색이 된 데이터 이미지를 이용하여 FPM으로 이미지를 가져온 후 딥러닝의 하나인 Generative Adversarial Network (GAN) 알고리즘을 이용하여 Phase image와 Amplitude image를 학습한 후 염색이 되지 않은 Phase image를 디지털 염색하는 방법이다. FPM은 같은 영역대 파장의 빛을 서로 다른 입사 각도에서 비추어 모델링을 통해 변환해 Image data를 얻는다[14-15]. 획득된 데이터 값을 이용하여 Image data를 추출할 수 있는데 이를 결합함으로써 하나의 확장된 이미지 및 광역 이미지를 만들 수 있다[16].

FPM으로 나온 데이터 값은 위상차 혹은 진폭을 이용한 매개변수에 따라 이미지를 추출한다. Phase image는 각도에 대한 매개변수를 이용해서 추출한다. Amplitude image는 진폭을 이용하여 데이터를 확보한다. 그리고 해당 값을 정규화하면 이미지를 얻을 수 있는데 정규화를 통해 식별가능한 이미지를 추출한다. Phase image는 형태에 대한 정보를 얻어오기 때문에 염색된 이미지와 염색되지 않은 이미지의 데이터가 서로 같다 그래서 완벽하게 위상이 같은 이미지를 얻을 수 있다.

학습 방법으로는 일반적인 GAN 알고리즘 방법과 비슷하게 먼저 Generator를 사용하지 않고 Discriminator가 Amplitude image를 정확하게 판독할 수 있을 때까지 정확성을 높인다. 다음으로 Discriminator를 멈추고 Generator를 학습시킨다. Generator는 Phase image를 가지고 Real Amplitude image와 비슷한 Fake Amplitude image를 만든다. 두 과정을 반복하여 최종적으로 Real Amplitude image의 특성이 반영된 결과를 추출하도록 유도하였다. 여기서 Generator에서 Phase image로 만든 Real Amplitude image와 가장 비슷한 Fake Amplitude image를 만들었다. Fig. 1.은 우리가 제안하는 Pair method의 다이어그램을 나타낸다.

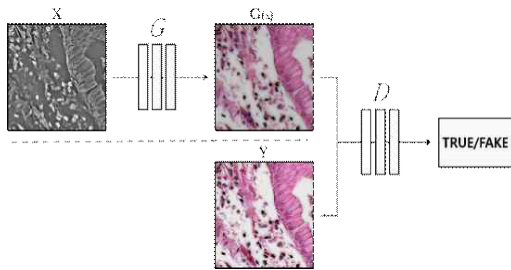


Fig. 1. Pair method 다이어그램

2.2 Unpair Method

Unpair Method는 일반적인 광학현미경으로 획득한 영상을 레이블 없이 unsupervised learning 기법을 사용하여 디지털 염색하는 방법이다. 앞서 시행한 phase에서 amplitude로의 변환처럼 1:1 pair model 학습 방법과는 다르게 X, Y 학습을 통한 X' 로 가기 위해서는 Unpair 방법을 사용해야 한다. 그러나 model

collapse 문제로 기존 생성자에서 다양한 이미지를 생성하지 못하고 비슷한 이미지를 생성하는 에러로 인해서 학습이 불가능하다. 그래서 기존 방식을 보완한 네트워크의 개념을 도입하여 X가 G encoder와 G decoder를 통과하면 X가 아닌 X' 를 생성하여 학습시 X와 X' 의 차이인 Cycle-consistency loss를 포함하여 학습을 진행한다. Adversarial Loss인 기존 GAN의 손실값과 Cycle-consistency loss의 값을 합쳐서 네트워크를 구성하면 Cycle-consistency loss 앞의 가중치를 GAN loss와의 상대적 중요도에 따라서 구성하여 전체적으로 구성된 objective function을 만들고 모델링이 형성된다. 해당 모델링은 pair가 되지 않고도 기존에 염색되지 않은 세포나 손상된 세포도 염색 후의 데이터 없이 기존 학습 모델링으로 학습가능하다.

III. 연구 결과

3.1 Pair Method Results

첫 번째 알고리즘인 Pair Method를 증명하기 위해 FPM에서 얻은 데이터를 바탕으로 데이터셋을 구축하여 이용했다. 데이터셋을 GAN 모델에 트레이닝 데이터로 만들어 모델링을 형성했다. Phase image, Amplitude image 각각 1,141개 이미지로 총 2,282개의 이미지와 Epoch는 500번으로 반복 학습하였다.

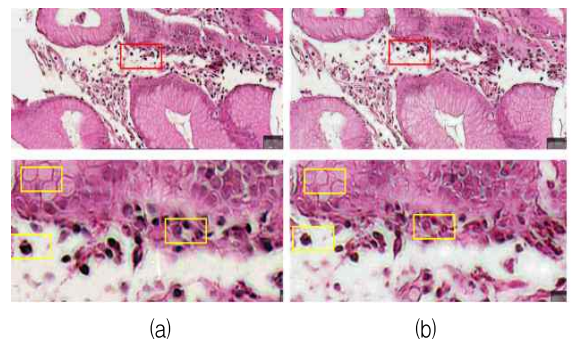


Fig. 2. (a) 실제 H&E 염색 영상, (b) Pair 방법으로 생성된 디지털염색 영상

Fig. 2.(a)는 염색된 슬라이드 실제 이미지이고 Fig. 2(b)는 염색되지 않은 슬라이드를 GAN 모델을 이용해 디지털 염색한 결과이다. 세포핵이나 세포질 부분이

육안으로 확인했을 때 거의 유사한 것을 확인하였고 상피조직, 적혈구 등도 이상 없이 구분하여 염색한 것을 확인하였다.

3.2 Frechet Inception Distance (FID)

디지털 염색된 이미지와 실제 염색된 이미지가 얼마나 같은지 확인하기 위해 Frechet Inception Distance (FID) 지표를 사용했다. FID 지표는 실제 이미지와 생성된 이미지에 대해 얼마나 유사한지 벡터 사이의 거리를 계산한다. Fig. 3은 pair methods를 통해 디지털 염색된 영상의 정확도를 보여주고 있다. 여기서 FID는 디지털 염색된 영상과 실제 H&E 염색된 영상 사이의 오차를 보여주고 있다.

AICD는 제안된 알고리즘의 이름으로 FID 는 평균 약 100 distance를 나타내고 있다. Gaussian1과 Gaussian2는 원본 영상에 임의로 Gaussian 잡음을 추가하고 밴드패스 필터로 필터링하면서 AICD 알고리즘으로 디지털 염색한 결과이다. Gaussian1 FID 는 약 50 distance, 그리고 Gaussian2 FID 는 약 102 distance 정도이다. 이를 통해 원본 영상에 Gaussian 잡음이 추가 되었을 때와 비교해서 제안된 AICD 알고리즘이 잡음 등에 robust 한 것을 알 수 있다.

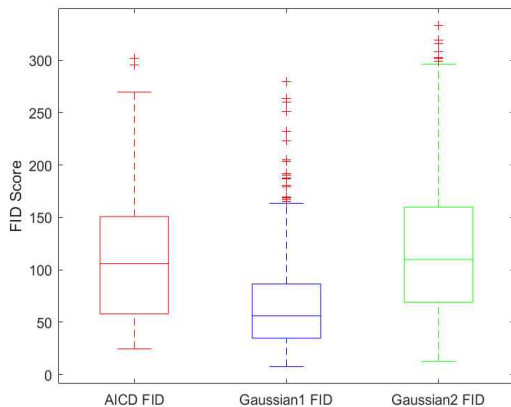


Fig. 3. 실제 영상과 디지털 염색(AICD), Gaussian noise#1 추가 후 디지털 염색(Gaussian1), Gaussian noise#2 추가 후 디지털 염색(Gaussian2)시 FID 비교

3.3 Unpair Method results

두 번째 알고리즘인 unpair 알고리즘을 수행하기 위해서 광학 이미지의 염색 영상과 염색되지 않은 영상 각각 24,282건을 사용하였다. Epoch는 500번으로 반복 학습하였다. Fig. 4. 는 실제 H&E 염색 이미지와 unpair 알고리즘으로 생성된 디지털 염색 영상을 보여 주고 있다. 실제 염색 이미지와 디지털 염색 이미지를 비교해 보았을 때 이미지의 전체적인 형태는 확인되며 상피조직의 인식까지 가능한 것으로 확인되었다.

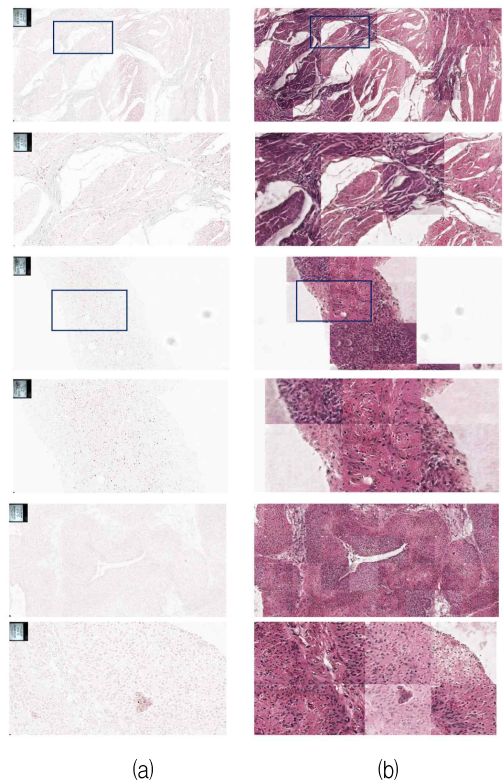


Fig. 4. (a) 실제 H&E 염색 영상, (b) Unpair 방법으로 생성된 디지털염색 영상

V. 결론 및 고찰

본 논문을 통해 두 가지의 디지털 염색 방법으로 염색을 진행해 보았다. Pair Method의 경우에는 FPM으로 촬영해야 하므로 데이터 확보가 어렵기는 하지만 pixel to pixel로 데이터 확보가 가능하므로 supervised learning이 가능하다. 디지털 염색 시 정확성 또한 높

다. Unpair Method에서는 unsupervised learning으로 비매칭 데이터 확보는 쉬우나 정형화 되어 있지 않아서 Pair Method에 비해 낮은 성능을 보여준다. 또한 디지털 염색 결과에 대한 정확한 측정 방법 역시 필요하다. 향후 두 가지 방법을 조합하여 상호 보완한다면 현재 연구 방법보다 우수한 결과를 나타낼 것으로 기대된다.

ACKNOWLEDGMENTS

This research is financial supported by “The digital pathology-based AI analysis solution development project” through the Ministry of Health and Welfare, Republic of Korea (HI21C0977) and Korea Basic Science Institute (National research Facilities and Equipment Center) grant funded by the Ministry of Education.(grant No.2020R1A6C101B189).

REFERENCES

- [1] D. Mayerich, M. J. Walsh, A. Kadjacsy-Balla, P. S. Ray, S. M. Hewitt, and R. Bhargava, “Stain-less staining for computed histopathology,” *Technology*, vol. 3, no. 01, pp. 27-31, 2015
- [2] F. Yi, J. Huang, L. Yang, Y. Xie, and G. Xiao, “Automatic extraction of cell nuclei from H&E-stained histopathological images,” *Journal of Medical Imaging*, vol. 4, no. 2, p. 027502, 2017.
- [3] K. de Haan et al., “Deep learning-based transformation of H&E stained tissues into special stains,” *Nature communications*, vol. 12, no. 1, pp. 1-13, 2021.
- [4] G. Avwioro, “Histochemical uses of haematoxylina review,” *Jpcs*, vol. 1, no. 5, pp. 24-34, 2011.
- [5] P. Marshall and R. Horobin, “The oxidation products of haematoxylin and their role in biological staining,” *The Histochemical Journal*, vol. 4, no. 6, pp. 493-503, 1972.
- [6] G. Clark, “Comparison of various oxidants for alum hematoxylin,” *Stain Technology*, vol. 49, no. 4, pp. 225-227, 1974
- [7] P. M. Kulkarni et al., “Deep Learning Based on Standard H&E Images of Primary Melanoma Tumors Identifies Patients at Risk for Visceral Recurrence and DeathDeep Learning-based Prognostic Biomarker for Melanoma,” *Clinical Cancer Research*, vol. 26, no. 5, pp. 1126-1134, 2020.
- [8] K. Kim and J. Guck, “The relative densities of cytoplasm and nuclear compartments are robust against strong perturbation,” *Biophysical Journal*, vol. 119, no. 10, pp. 1946-1957, 2020.
- [9] M. Mir, B. Bhaduri, R. Wang, R. Zhu, and G. Popescu, “Quantitative phase imaging,” *Progress in optics*, vol. 57, no. 133-37, p. 217, 2012
- [10] K. He, G. Gkioxari, P. Dollár, and R. Girshick, “Mask r-cnn,” in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 2961-2969.
- [11] F. Zernike, “Phase contrast, a new method for the microscopic observation of transparent objects part II,” *Physica*, vol. 9, no. 10, pp. 974-986, 1942.vol. 3, no. 01, pp. 27-31, 2015
- [12] J. Hosang, R. Benenson, P. Dollár, and B. Schiele, “What makes for effective detection proposals?,” *IEEE transactions on pattern analysis and machine intelligence*, vol. 38, no. 4, pp. 814-830, 2015.
- [13] G. Zheng, R. Horstmeyer, and C. Yang, “Wide-field, high-resolution Fourier ptychographic microscopy,” *Nature photonics*, vol. 7, no. 9, pp. 739-745, 2013.
- [14] M. Levoy, R. Ng, A. Adams, M. Footer, and M. Horowitz, “Light field microscopy,” in *ACM SIGGRAPH 2006 Papers*, 2006, pp. 924-934.
- [15] M. Levoy, Z. Zhang, and I. McDowall, “Recording and controlling the 4D light field in a microscope using microlens arrays,” *Journal of microscopy*, vol. 235, no. 2, pp. 144-162, 2009
- [16] B. R. Masters, “Quantitative phase imaging of cells and tissue,” *Journal of Biomedical Optics*, vol. 17, no. 2,sp. 029901, 2012.

저자소개

황 석 민 (Seok-Min Hwang)



2020년 6월 : 계명대학교
의용공학과(공학석사)
2023년 2월~현재 : 계명대학교
의용공학과 박사과정
관심분야 : Artificial Intelligence,
Data Analytics, 머신러닝, 지능화
시스템

조 남 훈 (Nam Hoon Cho)



1990년 2월 : 연세대학교
의과대학(석사)
1995년 2월 : 연세대학교
의과대학(박사)
2005년 2월~현재 :
연세대학교 의과대학 교수
관심분야:병리학, Genotype,
Immunohistochemistry,

정 연 우 (Yeun-Woo Jung)



2022년 ~현재 : 계명대학교
의용공학과(석사)
관심분야 : Artificial Intelligenc
데이터 처리, 머신러닝

이 종 하 (Jong-Ha Lee)



2011년 8월 :Temple University,
전자공학과(공학박사)
2012년 9월 : 삼성전자 종합기술원
2012년 3월~현재 :
계명대학교 의예과/의용공학과
교수
관심분야 : Medical Image,
Biosignal, Artificial Intelligence, Data Analytics

김 동 범 (Dong-Bum Kim)



2023년 2월 : 계명대학교
의용공학과(학사)
2023년 2월~현재 : 계명대학교
의용공학과(석사)
관심분야 : Artificial Intelligence,
Deep Neunal Network, 센서,
데이터 처리

이 승 아 (Seung Ah Lee)



2009년 2월 : 서울대학교
전자공학과(공학석사)
2014년 2월 : Caltech
전자공학과(공학박사)
2018년 2월~현재 :
연세대학교 전자공학과 교수
관심분야 : 바이오헬스, 인공지능
HW/SW, 시스템 반도체, 에너지-네트워크 통신