

금나노입자와 유전조작된 FD-바이러스 상호작용 분석에 관한 딥러닝 적용

송문혁¹², 신은섭¹², 이 용 오², 조 준², 레온 아벨만³, 누리예 지펠³
경희대학교 컴퓨터공학과¹, KIST 유럽연구소 스마트융합사업단², 바이오센서/물질연구단³
{[firefinger07](mailto:firefinger07@gmail.com), [kairos9603](mailto:kairos9603@gmail.com)}@gmail.com,
{[yongoh.lee](mailto:yongoh.lee@kist-europe.de), [j.jo](mailto:j.jo@kist-europe.de), [l.labelmann](mailto:l.labelmann@kist-europe.de), [n.korkmaz](mailto:n.korkmaz@kist-europe.de)}@kist-europe.de

Deep Learning Application of AuNP Interaction Analysis with Genetically Engineered FD-Viruses

Moon-Hyeok Song¹², Eun-Seop Shin¹², Yong Oh Lee², Jun Jo², Leon Abelmann³,
and Nuriye Zirpel³

Dept. of Computer Engineering in KyungHee University¹,
Smart Convergence Group² and Bio Sensor and Material Group³
in Korea Institute of Science and Technology Europe

요 약

본 논문은 안정적 담체시스템 개발에 있어 중요한 바이러스와 금나노입자의 결합 분석을 함에 있어 딥러닝 기술 중 시맨틱 분할(Semantic Segmentation)과 CNN(Convolution Neural Network) 분류자를 결합한 분석 모델 기법에 대해 제안한다. 제안 기법은 1) 시맨틱 분할을 통해 실험이미지로부터 바이러스 펩타이드와 금나노입자 모형을 분할하고 2) CNN 분류자를 통해 두 모형간의 결합 상태를 판별하는 모델이다. 제안된 모델은 94.2%의 정확도로 결합상태를 분석하며 기존 방식 대비 데이터 전처리 복잡도가 낮은 장점이 있다.

1. 서 론

무독성, 비용효율적인 안정적 담체시스템은 신약개발 및 암 치료 응용분야에서 중요한 기술이다. 금나노입자(AuNP, Gold Nanoparticle)는 화학적 합성을 통해 기하학적 특성을 가질 수 있지만 표적화 가능 물질을 얻기위한 입자의 기능화 연구는 현재 진행 중이다. 현재까지 항체, 비오틴분자, 화학반응과정을 사용하여 세포표적부위로 기능화하는 방법[1], 금(Au)에 선택적으로 결합하는 펩타이드 (Peptide)를 이용해 세포표적서열화하는 방법[2]이 연구되었다. [2]의 방식의 다양한 적용에 있어 바이오패닝(Biopanning)에 의한 펩타이드를 분리한 후 펩타이드 기능성에 대한 분석이 필요하며, 기능성화된 펩타이드(FD-바이러스)와 금나노입자 간 합성 조건에 대한 분석이 필요하다. 합성 조건에 대해 몬테카를로스 시뮬레이션을 통한 분석[3]이 있었지만, 이 연구는 엔트로피 변화가 반영이 안 되었고, 분자역학(Molecule Dynamics) 시뮬레이션의 한계인 원자레벨의 상호작용만을 고려한 문제가 있다.

본 논문에서는 펩타이드와 금나노입자의 결합을 분석하는데 있어 새로운 실험방법을 소개하고 딥러닝을 활용한 이미지 분석방법을 제안한다. 이 실험방식에서는 유체반응 실험장치(Fluid Reaction Chamber, 그림 1)를 통해 3D 프린트된 크기를 확대시킨 펩타이드와 금속입자 모형을 자유운동시키면서, 자성을 입힌

모형들간의 결합형태를 분석한다. 이 실험방식은 결합형태에 대한 가시화 및 실험조건 다양화(예: 금속입자 농도, 펩타이드 수, 외부온도 등)를 손쉽게 제어할 수 있다는 장점이 있다. 모형간 결합형태는 두대의 카메라를 통해 동영상으로 저장된다. 저장된 동영상을 분석하는데 있어 기존에는 전문가의 육안 혹은 복잡한 신호처리방식을 통해 이루어졌으나 딥러닝을 통한 이미지 분석을 한다면 정확성 및 효율성을 높힐 수 있다.

본 논문의 구성을 다음과 같다. 2장에서 3D 프린팅을 활용한 모형 디자인에 대한 설명하고 3장에서 모형결합 분석을 위한 딥러닝 적용 방법에대한 설명을 한다. 4장에서는 결론 및 향후 발전 방향에 대한 논의를 한다.



그림 1. 유체반응 실험장치

2. 3D 프린팅을 활용한 모형 디자인

유체반응 실험장치를 활용하기 위해서는 크기가 확대된 펩타이드와 금나노입자의 모형이 필요하다. 본 실험에서는 결합에 관여하는 펩타이드의 헤드부분의 일부를 모형화하였다(그림2.(a)). 실험에 사용된 펩타이드 모형과 금나노입자 모형의 디자인 정보는 그림2.(b)에서 확인할 수 있다. 실제 결합환경을 모사하기 위해 기하학적으로 계산된 수치를 이용하여 설계하였으며 모형을 3D 프린팅하는데 있어 방수가 가능한 재료를 사용하였다. 결합 실험을 위해 영구자석을 3D 프린팅된 모형에 부착하였으며(그림2.(b)의 주황색 부분) 유도되는 분자간 자력[4]을 고려해 자성의 크기 및 위치를 정했다.

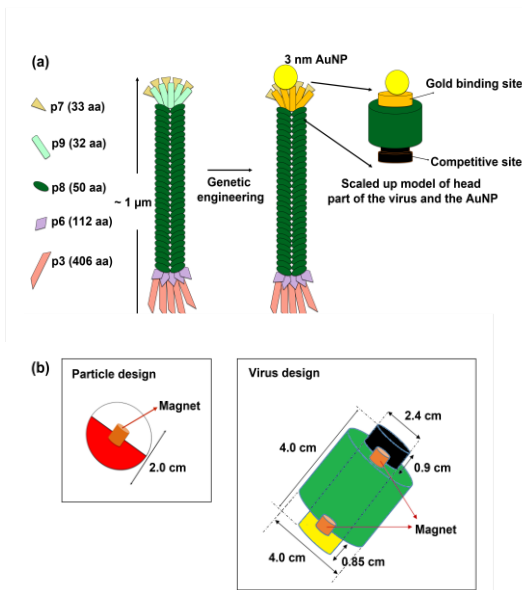


그림 2. 3D 프린팅된 펩타이드 및 금나노입자 모형

3. 결합모형 분석을 위한 딥러닝

실험 환경은 유체반응 실험장치 (그림1)에 3D 프린팅된 펩타이드와 금나노입자 모형 (그림2)을 넣어 유체운동 환경에서 진행하였다. 카메라 2대를 사용해 실험을 녹화했으며, 정면과 우측 이미지 RGB 1280(너비) x 1024 (높이) 크기의 3,193장의 실험 데이터를 추출했다. CNN에 적용함에 있어 데이터 크기가 큰 편이라 256x256 크기로 축소 변환하는 전처리를 진행했다. 변환한 데이터를 Cross-Validation 하기 위해 7 : 3 (Train : Test) 비율로 나누어 실험을 진행하였다.

실험에서 얻은 데이터 분석을 위해서 2단계로 구성된 딥러닝 모델을 제안한다. (그림3) 1단계는 시멘틱 분할 (Semantic Segmentation, [5])을 응용한 객체검출 단계로 주어진 이미지에서 펩타이드와 금나노입자 모형을 분할한다. 시멘틱 분할은 Convolution Nets[6]와 Deconvolution Nets[7]로 구성된다. 2단계는

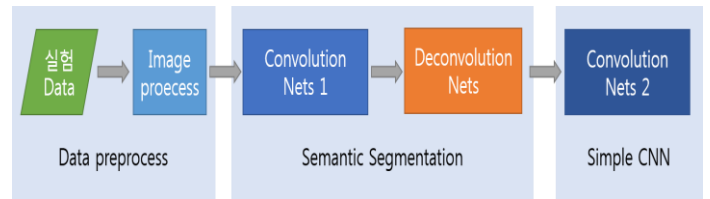


그림 3. 데이터 전처리 및 2단계 딥러닝 구조



그림 4. Semantic Segmentation의 Convolution Nets(좌)과 Deconvolution Nets(우)

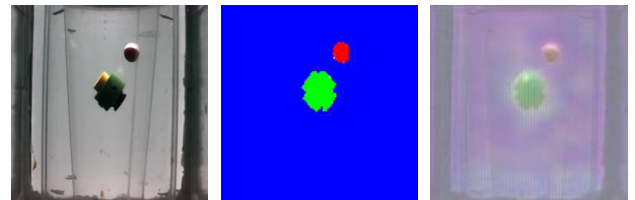


그림 5. 시멘틱 분할 실험 데이터(원본, 마스킹, 출력)

결합상태를 구분하는 분류자 Simple CNN이다. 시멘틱 분할된 정면과 우측 이미지를 각 Simple CNN을 지나 나온 결과를 통합해 금나노입자-펩타이드 결합여부를 판단한다.

시멘틱 분할은 기존 이미지 분할에서 쓰이는 경계상자가 아닌 픽셀 단위로 분할하는 방식으로 본 논문의 실험 결과의 정확성을 높여준다. 시멘틱 분할의 학습을 위해 300장 데이터의 객체를 마스킹(그림5 중간)하였다. 마스킹된 데이터들은 Convolution Net1(그림4,좌측)에 입력으로 사용되며, Convolution Net1의 출력이 Deconvolution Nets(그림4,우측)의 입력으로 사용되며 시멘틱 분할 모델이 학습된다. Batch Normalization(BN)[8]을 적용시켜 Banishing Gradient를 방지했으며, Activation Function은 Relu를 적용했다. 그림5 우측은 시멘틱 분할을 통해 펩타이드와 금나노입자 모형이 추출된 예이다.

시멘틱 분할 모델 학습이 완료되면 Simple CNN(그림 6)을 통해 금나노입자-펩타이드 결합여부를 구분한다. Simple CNN는 2종의 입력 데이터(정면과 우측 이미지) 데이터를 입력으로 사용하며, 각각 데이터는 Convolution Layers 4개와 Max Pooling 2개를 거친 후 통합되어 추가 Convolution Layer 2개, Max Pooling 1개를 통과 후 사전에 결합상태에 대한 라벨링 결과와 함께 학습된다.

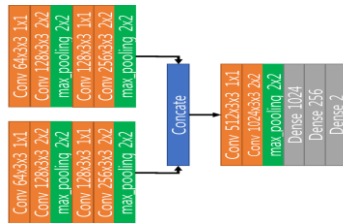


그림 6. Simple CNN

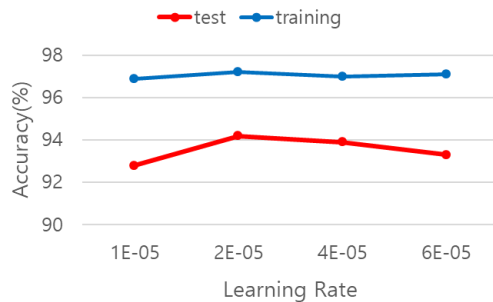


그림 7. Learning Rate 변화에 따른 정확도

본 논문의 실험을 진행한 결과 시멘틱 분할 단계의 Learning Rate = $2E-5$ 일 때, Train Accuracy = 97.2%, Test Accuracy = 94.2%가 나왔다. (그림 7)

4. 결론 및 향후 연구

서론에서 언급되었듯이, 생물학적 응용을 위한 나노크기의 담체 개발은 매우 중요하며, 이를 위해서는 금나노입자의 기본 결합 메커니즘의 이해가 필요하다. 이를 위해 우선 금나노입자와 펩티드 간의 상호 작용에 대한 연구가 탐구되어야 한다.

본 논문에서는 [9]에서 제안한 펩타이드-금나노입자 결합 분석 방법은 복잡한 신호처리 방식을 이용하는 한계가 있다. 이를 해결하기 위해 전처리 과정을 단순화시키기 시멘틱 분할 및 결합상태 분류하는 2단계 딥러닝 방식을 제안하였다.

본 논문의 실험결과 94.2%의 정확도를 보여 기존 복잡한 신호처리 방식을 대체할 수 있다. 나아가 복수의 펩타이드-금나노입자 등 다양한 조건에 대한 결합 분석을 할 때 전처리 단계의 복잡도를 줄여서 실험의 효율성을 높힐 것으로 예상된다.

향후에는 현재 이미지의 특징 추출 및 Segmentation을 담당하는 FCN 네트워크를 Neural Net 모델과 CRF과 결합되고 Dilated Convolutions[10] 기술이 적용된 DeepNetv3[11]로 바뀌어서 모델 정확도 성능 향상을 높힐 계획이다.

* 본 연구는 과학기술정보통신부 및 정보통신기술진흥센터의 SW중심대학지원사업(2017-0-00093)과 2018년 한국과학기술연구원 유럽연구소 기관 고유사업의 지원을 받아 수행된 연구임. (과제번호11807, Deep Learning Application of AuNP Interaction Analysis with Genetically Engineered FD-Viruses)

참 고 문 헌

- [1] D. A. Giljohann et al., "Gold nanoparticles for biology and medicine", in *Angewandte Chemie International Edition*, vol. 49, 2010.
- [2] A. Kumar et al., "Gold nanoparticles functionalized with therapeutic and targeted peptides for cancer treatment", in *Biomaterials*, vol. 33, 2012.
- [3] M. Chiappini et al., "Phase behaviour in complementary DNA-coated gold nanoparticles and fd-viruses mixtures: a numerical study", in *The European Physical Journal*, vol. 40, 2017.
- [4] T. Hageman et al., "Macroscopic equivalence for microscopic motion in a turbulence driven three-dimensional self-assembly reactor", in *arXiv preprint arXiv:1709.04978*, 2017.
- [5] Jonathan Long et al., "Fully Convolution Networks for Semantic Segmentation," in *arXiv preprint arXiv:1411.4038*, 2015.
- [6] Alex Krizhevsky et al., "ImageNet Classification with Deep Convolutional Neural Networks", in *NIPS*, 2012.
- [7] Hyeonwoo Noh et al., "Learning Deconvolution Network for Semantic Segmentation", in *arXiv preprint arXiv:1505.04366*, 2015.
- [8] Sergey Ioffe et al., "Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift", in *arXiv preprint 1502.03167*, 2015.
- [9] Haebom Lee et al, "Deep Learning Analysis of Binding Behavior of Virus Displayed Peptides to AuNPs", in *12th International Conference on Practical Applications of Computational Biology & Bioinformatics*, 2018
- [10] Fisher Yu, et al., "Multi-Scale Context Aggregation by Dilated Convolutions", in *arXiv preprint arXiv:1511.07122*, 2015.
- [11] Liang-Chieh Chen et al, "Rethinking Atrous Convolution for Semantic Image Segmentation", in *arXiv preprint arXiv:1706.05587*, 2017.