

<주요 Q&A>

Efficiently and Effectively developing Medical Imaging AI Systems

- Q1. Ai로 질병을 진단하니 혁신이네요 그런데 이렇게 진단을 받고 나면. 최종적으로 사람의사가 검수를 하나요 그 자체만으로 의학적 진단으로 검증이 가능한가요? 의학계에서 반발이 심할 텐데 협의는 되고 있는지 궁금합니다

의사들과 협업을 통해서 기능 검증과 개선, 논문작업도 진행하고 있습니다.

- Q2. AI가 진단한 이후에는 의사의 소견이 얼마나 포함되는지 궁금합니다. 의학영상 외에도 의학의 다른 영역에 AI가 어떻게 사용되는지 궁금합니다.

모델은 현재 의사의 최종 진단을 도와주는 용도로 개발 중이고, 헬스케어 관련 모델을 주로 개발하고 있습니다.

- Q3. 이런 진단이 원격으로 가능한 개념인 것인가요?

네, 클라우드를 통한 원격 개념입니다.

- Q4. 메타버스+웨어러블 의료장비+AI를 통해 원격의료가 일상화 되는 날이 언제쯤 구현될 수 있을까요?

조만간 현실화 되길 기대해봅니다.

- Q5. 인종적 차이도 고려한 라벨링이 가능 하려면 어떤 보완이 필요할까요?

좋은 질문이십니다. 자세한 설명은 어렵지만 데이터, 모델 관점에서 모두 개선이 가능한 방법이라고 봅니다.

Q6. 오토라벨링도 오차가 있을 것 같은데 어느 수준인가요?

모든 데이터셋으로 생성한 레이블링과 같은 정도의 오차로 개발하고 있습니다.

Q7. 의료 영상 분석은 개인정보규제로 국내에선 연구 목적으로만 사용할 수 있다고 알고 있는데, 해외에서는 영리목적으로 데이터셋 사용이 가능한가요?

연구용으로는 오픈 데이터셋을 주로 참고하고, 영리목적 모델인 경우 고객사의 데이터를 받아서 모델을 개발하고 데이터를 폐기하는 과정으로 진행하고 있습니다.

Q8. I think eXplainable AI(XAI) is very important in health care, especially in medical image recognition. Is SDS applying XAI?

네, 설명가능하고 검증 가능한 모델로서 개발 중에 있습니다.

Q9. 치과에서는 물고 있는 것으로 구강 내 산성도나 균 분포도를 분석하고 진단하는 기기가 있다고 들었어요 이외에도 실용에 가깝게 만들어진 사례가 있나요?

치과 관련 모델도 현재 고객사와 협업중에 있습니다.

Q10. 종양이 피사되었거나 종양이 계속 자라나는 경우의 오토라벨링으로 잘 구분할 수 있을지 궁금합니다.

이 부분도 데이터 레이블 작업이 아주 중요합니다. 그 후에 모델에서 어느 정도 보완가능하고요.

네, 데이터가 잘 준비되면 가능할 것 같습니다.

Q12. 클라우드의 단점 보완은 어느 정도 일까요?

클라우드 데이터 보안 부분은 프라이버시 문제로 미국에서도 법적으로 규제를 하고 있습니다.

Q13. 질병 기반 의료 이미지를 기반으로 질병을 탐지하고 진단하는 데 인공지능 기술을 적용하고 활용하는 경우 핵심적으로 검토하고 테스트해야 할 요소들에 대해서 문의 드립니다.

가장 중요한 부분은 모델의 신뢰성과 데이터 레이블의 정확도라고 생각합니다. 데이터의 정확도가 보장된 상황에서 모델이 제한하는 결과값의 올바른 해석이 가능합니다.

Q14. AutoLabel을 하기 위한 최소한의 레이블링 작업은 어느 정도일까요?

데이터셋마다 다르지만 10%정도 레이블링으로도 가능합니다.

Q15. 데이터 셋마다 다르다는 것은 데이터의 종류에 따라서 다르다는 것이지요? 예를 들어, 의료 이미지의 경우에는 10% 정도면 가능하지만, 다른 일반적/일상적 이미지는 다를 수 있다는 것일까요?

네, 데이터 종류에 따라서 난이도 차이가 납니다. 가령, 클래스 개수가 10개인 데이터셋과 100개인 데이터셋인 경우 보통 클래스 10개인 데이터셋이 더 적은 레이블링 작업을 요구합니다.

Q16. 이런 기술의 발전은 의사들의 반발이 클 것 같은데 그래서 법률의 제정도 어려울 것 같은데 현재 헬스분양에서 Ai를 사용할 수 있는 기반 법이 마련되어 있는지요?

좋은 질문입니다. 이진 협업을 통해서 함께 풀어나가야 할 문제라고 생각합니다.

Q17. 휴먼 라벨링팀이 active learning model로 데이터를 신속하게 보내고 효율적으로 처리할 수 있는 방법에 대해서 질문 드립니다. 이 경우 고려해야 할 중요한 요소들에 대해서 질문 드립니다.

네, 잘 설계된 모델, 그리고 그 모델을 통해 데이터셋의 Uncertainty 값을 잘 측정할 수 있어야 하고, 휴먼 레이블의 정확도도 매우 중요하게 작용합니다.

Q18. 강연하시는 논문 제목을 알려주실 수 있나요?

<https://arxiv.org/abs/2105.14559> 조만간 최종 본 업데이트 예정입니다.

Q19. 의료기기 사용시 개인정보 등 보안문제에 관해 어떻게 해결하실지 궁금합니다.

개인정보는 익명데이터로 제공되고 모델 개발 시에만 사용되고, 개발 후 폐기됩니다.

Q20. 초음파, CT, MRI 등 수집된 영상 데이터를 AI 이미지 분석을 고도화하기 위한 방안은 어떻게 되는지요? AI 이미지 분석 모델링 검증과 수정은 어떻게 진행되는지요?

고객사와 현업사와의 협업을 통해서 함께 검증작업을 진행하고 있습니다.

Q21. 레이블이 중요한 요소인데 레이블링이 10%가 안되고 나머지는 예측할 수 있는 것은 모델링에 따라 다른 것인가요?

경험적으로 모델링 보다는 데이터자체가 가진 분포와 클래스 개수와 같은 문제의 난이도 요소가 더 큰 것 같습니다.

만큼 어떻게 보안관리를 할 수 있을지 궁금하네요.

클라우드의 기본 보안 절차를 따라가지만, 익명화된 개인정보와 같은 민감한 데이터는 암호화 과정을 거쳐서 데이터 사용과 폐기가 함께 이루어지고 있습니다.

Q23. baba가 의료영상에 특별히 성능이 높은 건가요? 아니면 이미지 세그멘테이션 성능이 높은 건가요?

의료영상, 일반적인 데이터 셋, 세그멘테이션 모두 우수한 성능을 보여주고 있습니다.

Q24. 이미지 객체를 자동으로 세분화해서 인식하는 건가요?

네, 객체에 대한 사전 정보 없이 이미지 객체를 추출하는 작업입니다.

Q25. 기존 의료장비들에 측정결과를 1차로 AI 가 유형별로 데이터를 보유하고, 환자가 와서 측정한 방사선 ct mri 등 이미지가 바로 시스템에 등록된 1차 자동처방 결과를 의사한테 보여주면 당일 검사한 결과를 보고 의사가 추가 노트 처방하는 구조가 되는 거겠죠?

네, 의사의 진단 시간/과정/노력 절감으로, 모두가 효율적이고 신속한 의료 혜택을 받을 수 있도록 노력하고 있습니다.

Q26. 앞서 다뤄주신 임베디드 AI의 중요성이 이번 의료도메인 주제와도 연결되는 지점이 있을 것 같네요, 가령 클라우드에 의료정보를 업로드하여 활용되는 것이 절대 완벽한 보안을 보장할 수 없을 것 같기도 하고, 임베디드 AI를 통한 오프라인 러닝이 가능하고 고효율의 학습이 단말 내에서 가능하다면 보다 안전한 의료데이터 관리와 활용이 가능할 까요?

네, 좋은 제안이라고 생각합니다. 고객사와의 협업에서도 비슷한 경우가 있었습니다.

Q27. medical imaging AI system을 병원에서 사용하는 경우 사용상 유의해야 할

사항들에 대해서 질문 드립니다.

의사의 진단을 도와주는 용도로 개발 중이므로 맹신은 금물이고, 개인 정보 데이터 처리에 좀 더 유의해야 합니다.

Q28. 주로 뇌, 혈관, 종양 등과 관련한 내용들이 많은 것 같은데 뼈나 눈 관련 데이터를 사용했을 때 성능이 얼마나 잘 나올지 궁금합니다.

X-Ray 이미지, 안구 이미지 등, 각기 다른 데이터셋에 대해서도 유사한 성능을 보여주고 있습니다만, 일반화된 척도로 성능을 설명하기는 어렵습니다.

Q29. 병원에서 의학 데이터 셋에 대한 모델링을 최적으로 할 수 있는 방법에 대해서 질문 드립니다 이 경우 핵심적으로 테스트해야 할 요소들에 대해서 질문 드립니다.

데이터와 문제에 따라서 적용하는 모델이 각기 달라서 일반화된 방법은 존재하지 않습니다. 가령, 2D 이미지인지 혹은 3D 이미지 인지에 따라 모델의 구조가 달라지고, X-Ray/CT/MRI 중 어떤 이미지 인지에 따라서 적용하는 방법이 각기 달라서 모델링 하는 사람의 노하우가 중요하게 작용합니다.