

유방암 진단을 위한 PCA 알고리즘 구현

충북대학교 의생명과학경영융합대학원 석사과정 김귀숙
충북대학교 의생명과학경영융합대학원 교수 손호선
충북대학교 의생명과학경영융합대학원 교수 류근호

최근 발표된 '2013년 한국여성 유방암 백서'에 따르면, 2011년 전체 유방암 환자 중 폐경후 여성(50대 이상)의 유방암 비율이 51.3%를 기록, 최초로 폐경전 여성의 발병률을 역전했다. 유방암 환자의 중간나이도 2000년 46세에 비해 4세 증가한 50세를 기록했다. 이러한 현상은 한국인 유방암 발병률 증가 및 전체적으로 연령이 높아지는 양상을 보이며, 발병 및 치료 모두에서 뚜렷하게 서구화 추세를 보이고 있음을 알 수 있다. 유방암은 조기 치료에 따른 환자의 생존률이 높은 암종 하나로 유방암의 진단은 매우 중요하다.

본 연구에서는 유방암의 진단을 위해 직접 구현한 PCA 알고리즘을 이용하여 속성을 선택하는 방법을 제시하고자 한다. 실험에 사용한 데이터는 UCI에서 제공하는 유방암 데이터이다. 속성 선택 기법으로는 직접 구현한 PCA 알고리즘을 이용하였다. 실험에서는 기존의 PCA 알고리즘을 이용한 분석 프로그램을 통한 데이터 속성 선택과 구현한 PCA 알고리즘의 속성 선택 결과를 비교 분석하여 구현한 알고리즘의 정확도 및 타당성을 판단한다. 다음으로 선택한 속성들을 여러 분류 기법을 이용하여 분류하여 그 결과를 비교한다. 기존의 분석 프로그램 내의 PCA 알고리즘과 직접 구현한 PCA 알고리즘을 비교한 결과, 같은 결과의 고유값과 주성분을 얻을 수 있었다. 따라서 우리가 실제 구현한 PCA 알고리즘이 정확하고 그 결과를 신뢰할 수 있음을 알 수 있었다. 이 연구를 통해 암을 보다 정확히 진단하기 위한 유의한 요인을 찾는 데 필요한 PCA 알고리즘을 쉽게 활용할 수 있고, 이를 활용하여 추후 환자의 암 예방 및 진단에 도움을 줄 수 있을 것으로 기대된다.

주제어 : 유방암 진단, PCA 알고리즘, 속성 선택
