

모호한 카테고리를 도입한 기저 세포암 검출기

*박아론, 민소희, 백성준, 나승유
전남대학교 전자컴퓨터공학부

e-mail : dodenet@daum.net, minsh@chonnam.ac.kr, tozero@chonnam.ac.kr,
syna@chonnam.ac.kr

A Basal Cell Carcinoma Classifier with an Ambiguous Category

*Aaron Park, Sohee Min, Seong-Joon Baek, Seung Yu Na
The School of Electronics and Computer Engineering
Chonnam National University

Abstract

According to the previous work, various well known methods including maximum *a posteriori* probability classifier (MAP) and multi layer perceptron networks classifier (MLP) showed competitive results. Since even the small errors often leads to a fatal result, we investigated the method that reduces classification error perfectly by screening out some ambiguous patterns. Those ambiguous patterns can be examined by routine biopsy. We incorporated an ambiguous category in MAP and MLP. Classification results involving 216 spectra gave 100% sensitivity for the case of MLP.

I. 서론

피부암은 세계에서 가장 일반적인 암 중 하나이다. 그 중 가장 일반적인 피부암은 기저 세포암 (Basal Cell Carcinoma: BCC)과 편평상피 세포암 (Squamous Cell Carcinoma)이다[1]. 이것들 중 BCC는 보다 빈번히 나타나는 피부 종양이고 주위의 암이 아닌 조직 (normal tissue: NOR)과 구별하기가 매우 어렵다. BCC는 보통 조직 샘플을 병리학적으로 검사하여 검출한다. 하지만 이 방법은 처리가 복잡하고 각 병리학자의 경험에 따른 주관적인 판단에 의지하며 때로는 정

확한 생체검사를 위해 과도한 조직을 필요로 한다[2]. 이전에 행한 연구에서 BCC 검출을 위해 MAP, MLP, PNN (Probabilistic Neural Networks), KNN (*k*-nearest neighbor)를 포함하는 여러 가지 분류 방법을 사용하였고 그 결과 분류 오류율이 4-5%정도였다 [3]. 때로는 이 오류가 치명적일 수 있기 때문에 본 연구에서는 모호한 카테고리를 도입하여 분류 오류율을 낮추는 방법을 제안하였다.

II. 실험방법

공초점 라만 분광기를 이용한 조직 샘플의 스펙트럼은 피부 표면으로부터 수직인 방향 30-40 μ m 간격으로 서로 다른 지점에서 측정하였다[2]. 이와 같은 방식으로 본 연구에서는 10명의 환자로부터 216개의 라만 스펙트럼을 얻었다. 측정 후 스펙트럼은 [-1, 1] 범위로 정규화하였다. 또한 불필요한 데이터를 자르기 위해 Half-Hanning Window를 사용하였고 차원을 줄이기 위해 Principal Component Analysis를 적용하였다[3]. 모호한 카테고리를 도입하기 위해 MAP와 MLP의 결정 규칙을 수정하였다. 그리고 첫 번째와 두 번째 주성분만을 이용하여 분류기의 성능을 분석하였다. MAP에서는 클래스 조건부 확률을 다변수 가우시안 확률 함수로 가정하고 $\omega_1, \omega_2, \omega_3$ 를 각각 BCC 클래스, NOR 클래스 그리고 모호한 클래스라 하였다. 이 때 g_i 를 판별함수, θ 를 경계값이라고 하면 MAP 결정 규칙은 다음과 같이 정했다.

* 본 연구는 산업자원부 지방기술혁신사업(RTI04-03-03) 지원으로 수행되었음.

$$\begin{cases} \omega_1 & \text{if } g_1(x) - g_2(x) \geq \theta, \\ \omega_2 & \text{if } g_2(x) - g_1(x) \geq \theta, \\ \omega_3 & \text{otherwise.} \end{cases}$$

MLP는 5개의 은닉 유닛과 하나의 출력 유닛으로 구성하고 NOR 클래스에 대해 -1, BCC 클래스에 대해 +1이 출력되도록 오류 역전파 알고리즘을 이용하여 훈련시켰다. MLP 네트워크의 출력값이 z일 때 MLP 결정 규칙은 다음과 같이 정했다.

$$\begin{cases} \omega_1 & \text{if } z > +\theta, \\ \omega_2 & \text{if } z < -\theta, \\ \omega_3 & \text{otherwise.} \end{cases}$$

그림 1에는 세 가지 클래스의 특징 벡터 분포를 나타내었다. x축과 y축은 첫 번째와 두 번째 주성분에 대응된다. 그림은 모호한 데이터들이 결정 경계의 근처에 있는 것을 분명하게 보여준다.

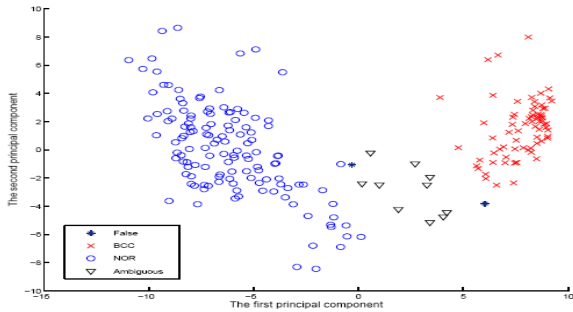


그림 1. MAP 분류기에 의한 BCC 카테고리, NOR 카테고리 및 모호한 카테고리의 패턴 분포.

III. 실험 결과

표 1에 모호한 카테고리를 적용하지 않았을 경우 대한 분류 결과를 요약하였다. 그 결과에 따르면 평균 specificity가 약 95.0%임을 볼 수 있다. MLP는 초기값에 따라 그 성능이 달라지므로 본 연구에서는 20회 실험의 결과를 평균하였다.

표 1. 모호한 카테고리를 적용하지 않았을 경우 대한 분류 결과 (*는 전문 병리학자의 결정임).

	MAP		MLP	
	BCC	NOR	BCC	NOR
BCC*	96.1	3.9	98.7	1.3
NOR*	3.6	96.4	6.4	93.6

표 2에는 모호한 클래스를 적용한 경우에 대한 실험 결과를 요약하였다. 이 때 사용한 경계값 θ 는 각각 6, 0.89이다. 모호한 클래스로 분류되는 패턴의 수는 테스트 패턴의 10%를 넘지 않도록 조절하였다. 그 결과에

따르면 평균 specificity가 약 99.2%이다.

표 2. 모호한 클래스를 적용한 경우에 대한 분류 결과 (*는 전문 병리학자의 결정임).

	MAP		MLP	
	BCC	NOR	BCC	NOR
BCC*	97.1	2.9	100	0
NOR*	0.8	99.2	0.8	99.2

하지만 여기에서 평균 specificity가 4% 이상 증가했다는 결과보다 MLP의 specificity가 100%라는 결과가 모호한 카테고리 도입의 의미를 잘 보여준다고 할 수 있다. MAP 또한 더 많은 수의 모호한 데이터를 허용하면 완벽하게 분류할 수 있다. 모호한 데이터를 증가 시킴에 따른 MAP의 분류율을 표 3에 나타내었다.

표 3. 모호한 데이터가 증가함에 따른 MAP의 분류 결과 (*는 전문 병리학자의 결정임).

	모호한 데이터의 백분율							
	9.2 %		10.5 %		12.7 %		16.2 %	
	BCC	NOR	BCC	NOR	BCC	NOR	BCC	NOR
BCC*	97.1	2.9	98.5	1.5	100	0	100	0
NOR*	0.8	99.2	0.8	99.2	0.8	99.2	0	100

이 실험 결과에서 볼 수 있듯이 모호한 카테고리를 도입하면 기존 분류기의 성능을 향상시킬 수 있으며 모호한 데이터 수를 보다 많이 허용하면 완벽한 분류가 가능하다는 것을 확인할 수 있다.

참고문헌

[1] A. Jijssen, T. C. B. Schut, F. Heule, P. J. Caspers, D. P. Hayes, M. H. Neumann and G. J. Puppels, "Discriminating Basal Cell Carcinoma from its Surrounding Tissue by Raman Spectroscopy," *Journal of Investigative Dermatology*, vol. 119, pp. 64-69, July 2002.

[2] J. Choi, J. Choo, H. Chung, D.-G. Gweon, J. Park, H. J. Kim, S. Park and C.-H. Oh, "Direct observation of spectral differences between normal and basal cell carcinoma(BCC) tissues using confocal Raman microscopy," *Biopolymers*, vol. 77, pp. 264-272, April 2005.

[3] S.-J. Baek and A. Park, "Basal Cell Carcinoma Detection by Classification of Confocal Raman Spectra," *to be appeared in ICIC 2006 Proceedings*, 2006.