

# 구순열 비변형의 객관적 평가를 위한 Neural Network의 적용

## Objective assessment of cleft lip nose deformity by neural network

박중훈\*, 김진태\*, 홍현기\*, 김수찬\*\*, 김덕원\*

Joong-Hoon Park\*, Jin-Tae Kim\*, Hyun-Ki Hong\*, Soo-Chan Kim\*\*, Deok-Won Kim\*

**Abstract** - Cleft palate is a congenital deformity condition with separation of the two sides of the lip resulting in nose deformity. Evaluation of surgical corrections and outcome assessments for nose deformity due to the cleft lip depends mainly on doctor's subjective judgment. An objective method for evaluation of the condition and surgical outcome of nose deformity due to the cleft palate is needed. This study aimed at objective assessment of a cleft palate nose deformity condition by analyzing the following parameters obtained from photographic images of a cleft palate patients: (1) angle difference between two nostril axes, (2) center of the nostril and distance between two centers, (3) overlapped area of two nostrils, and (4) the overlapped area ratio of the two nostrils. A regression equation of doctor's grades was obtained using the eight parameters. Three plastic surgeons gave us the grades for the each photographic image by 10 increments with maximum grade of 100. The average reproducibility of the grades given by the three plastic surgeons and the three laymen using the developed program was  $10.8 \pm 4.6\%$  and  $7.4 \pm 1.8\%$ , respectively. Kappa values representing the degree of consensus of the plastic surgeons and the three laymen were 0.43 and 0.83, respectively. Correlation coefficient of the grades evaluated by the surgeons and obtained by the neural network was 0.798. In conclusion, the developed neural network model provided us better reproducibility and much better consensus than doctor's subjective evaluation in addition to objectiveness and easy application.

**Key Words** : cleft palate, nose deformity, nostril area, neural network

### 1. 서론

구순열은 두경부에 발생하는 가장 흔한 선천성 기형 중의 하나로서 심한 코변형을 초래한다. 임상에서는 구순열 및 코변형의 일차적 수술 및 이차적인 교정을 시도한 후에 결과를 임상상의 주관적인 판단 하에 3 등급(good, fair, poor)이나 4 등급(mild, moderate, severe and very severe), 혹은 5등급(excellent, very good, good, satisfactory and poor)으로 나누어 평가하고 있다 [1].

구순열로 인한 코변형을 객관적으로 평가할 수 있다면 환자의 수술결과를 정량적으로 보여줌으로써 새로운 수술 기술의 평가 및 수술 예후 평가에 많은 도움을 줄 수 있을 것이다. 현재는 전문의가 앞에서 언급한 몇 가지 등급으로 평가하거나, 환자나 보호자의 판정과 비교 관찰하는 정도이다 [2-3]. 디지털 카메라에 의한 2D 영상이나 3D 스캐너를 이용한 입체 영상으로부터 구순열로 인한 코변형을 평가하고자 하는 시도가 있었으나 이러한 방법은 측정 파라미터가 너무 많고 3D 스캐너의 경우 장비가 고가인 단점이 있다 [4].

본 연구에서는 비공의 각도, 비공의 중심점, 비주를 기준으로 한 비공의 겹쳐진 면적 등의 인자를 대상으로 코 변형의 객관적

인 평가에 도움을 줄 수 있는 인자들을 통계적 분석법을 이용하여 찾았다. 이렇게 찾아낸 인자로부터 각각 숙련된 전문의 평가 점수와의 상관관계를 신경망을 이용하여 기존 임상상의 경험에 의한 주관적 분석 결과를 객관화하고, 재현성을 높이고자 하였다.

### 2. 코변형 평가를 위한 분석인자

구순열로 인한 코 변형 정도는 시각적인 관찰을 통해서 양측 비공(nostril)의 대칭성으로 판단할 수 있다. 양측 비공의 대칭성은 양측 비공 중에서 보다 정상에 가까운 측의 비공을 기준으로 반대 측 비공과의 비공 장축의 각도, 면적, 거리에 의하여 그림 1과 같이 분석 인자를 고려하였고, 이들 인자 중에서 대칭성에 영향을 많이 주는 인자가 무엇인가를 통계적으로 분석해 보았다. 그림 1에서 우측 비공을 정상, 좌측 비공을 비정상이라 가정하였을 때, 각각의 비공 장축의 각도, 면적, 거리는 다음과 같다.

S1 :정상측 비공의 면적

S1' :정상측 비공을 비주에 대하여 대칭으로 한 비공의 면적

S2 :비정상측 비공의 면적

S3 :S1'과 S2의 중첩 면적

C1 :정상측 비공의 중점

저자 소개

\* 연세대학교 의과대학 의공학학교실

\*\* 한경대학교 생물정보통신대학원

C1': 정상측 비공을 비주에 대하여 대칭으로 한 비공의 중점

C2: 비정상측 비공의 중점

$\theta 1$ : 정상측 비공 장축의 각도

$\theta 1'$ : 정상측 비공을 비주에 대하여 대칭으로 한 비공 장축의 각도

$\theta 2$ : 비정상측 비공 장축의 각도

D1: C1 과 C2 간의 거리

D2: C1' 과 C2 간의 거리

D3: C1 과 C1' 간의 거리

위에서 얻어진 값으로부터 정상 측과 비정상측의 비대칭성은 다음 분석 요인들을 고려할 수 있다.

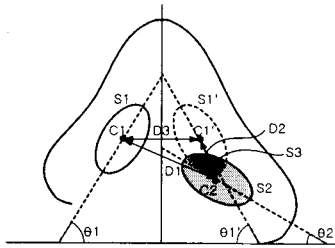


그림 1. 코 변형 평가를 위한 변수

### 1. 비공 장축의 대칭성

가.  $\theta 1 / \theta 2$ : 정상 측 비공의 각도와 비정상측 비공의 각도 비. 두 개의 각도 중 작은 값에 대한 큰 값의 비율로 계산한다. 정상일 경우  $\theta 1 = \theta 2$  가 되므로  $\theta 1 / \theta 2 = 1$ 에 가까워질수록 정상

나.  $|\theta 1 - \theta 2|$ : 정상 측 비공의 각도와 비정상측 비공의 각도 차이. 정상일 경우  $\theta 1 = \theta 2$ 를 만족하므로  $|\theta 1 - \theta 2| = 0$ 에 가까움

### 2. 비공 위치의 대칭성

영상의 배율에 따라 같은 거리일지라도 영상에서 계산된 거리의 차이가 개인과 영상 간에 있을 수 있으므로 이러한 영향을 줄이기 위해 각 거리의 차이보다 거리간의 비를 이용하였다. 비공 중점 간의 거리로 계산되는 비공 간의 거리에서 D2는 비변형 정도에 비례하는 인자이며, D3은 정상 대칭일 경우의 거리이다. D1/D3, D2/D3, D2/D1 세 가지 경우를 고려할 수 있다. 정상일수록, D1과 D3는 같아지고 D2는 0에 가까워지므로 D1/D3은 1, D2/D3 와 D2/D1은 0에 가까워진다.

### 3. 비공 면적의 대칭성

S2/S1, S3/S1, S3/S2, S3/(S1+S2), S3/(S1+S2-S3)의 다섯 가지의 경우를 고려해 보았다. 정상일수록, S1, S2, S3는 같아지므로 S3/(S1+S2)은 0.5에 접근하고, 나머지는 모두 1에 가까워진다.

## 3. 실험방법

세브란스병원 성형외과에서 코변형 교정수술을 받은 86명의

환자 영상을 분석하였다. 구순열 환자의 수술 경험이 많은 성형외과 전문의 3명으로부터 100점 단점의 10점 간격으로 사진을 하루 간격으로 3회씩 검토하여 평가점수를 기록하였다. 본 연구에서 개발한 컴퓨터 영상분석은 일반인 3명이 하루 간격으로 각각 3회씩 10개 ( $\theta 1 - \theta 2$ ,  $\theta 1 / \theta 2$ , D1/D3, D2/D3, D2/D1, S2/S1, S3/S1, S3/S2, S3/(S1+S2), S3/(S1+S2-S3))의 인자들을 분석하게 하였다. 이를 위해 LabVIEW 6.1(National Instruments, USA)과 NI-IMAQ Vision을 활용하여 프로그램 하였다.

일반인 3명이 영상에서 양쪽의 비공, 비공의 장축, 비주, 그리고 영상의 수평선과 같은 분석을 위한 정보를 수동으로 설정하면 앞의 모든 인자와 관련된 분석 결과는 자동적으로 얻을 수 있다. 획득된 인자로부터 통계 분석을 통해 재현성 및 임상과의 상관관계가 높은 인자를 추출하고, 신경망 모델을 이용하여 분석 인자와 임상 결과와의 상관관계를 학습을 통하여 추정하였다.

## 3. 결과

임상의 평가 점수에 대하여 임상간의 차이를 보기 위해 일치도 검사를 하였다 (표 1). 임상간의 일치도 검사는 Kappa 검정방법을 사용하였다. 통계값은 0.4미만: 낮은 일치도, 0.4~0.75: 보통, 0.75이상: 높은 일치도를 나타낸다. 각 임상간의 3회 측정결과는 3명의 임상의 결과는 보통 정도의 일치도를 보여 반복 측정에 의한 차이가 없음을 보였다. 또한, 표1과 같이 임상간의 일치도가 높지는 않지만 보통의 일치도를 갖는 것을 확인했다. 그러나 임상2의 값은 통계적으로 유의하지 않기 때문에 임상의 평가 점수에서 제외하고 2명의 임상의 평가 점수의 평균값으로 차후 분석을 하였다.

표 1. 임상간의 3회 테스트 평균값에 대한 통계량

Kappa Value	임상의1	임상의2	임상의3
임상의1	1	0.43	0.49
임상의2	0.43	1	0.36
임상의3	0.49	0.37	1

일반인 3명이 개발된 프로그램을 활용하여 분석한 10개의 파라미터 값은 ANOVA 검정을 통해 일반인의 시행횟수에 따른 차이를 검정하였다. 개인마다 3회 측정된 10개의 파라미터 간에는 모두 유의한 차이가 없었다( $p > 0.05$ ). 이로부터 컴퓨터를 활용한 분석은 수동으로 측정 인자를 설정하는 과정에서 개입될 수 있는 주관적 차이는 거의 없음을 확인할 수 있었고, 일반인의 테스트 횟수 간에도 영향을 주지 않음을 확인할 수 있었다. 컴퓨터 분석을 통해 얻은 10개의 파라미터 중에서 전문의의 판단 점수에 영향을 잘 반영하는 입력인자를 결정하기 위하여 파라미터의 평균값과 전문의 평균 평가 점수와의 상관관계 분석을 하여 어떤 파라미터가 전문의 평가점수에 많은 영향을 미치는지 알아보았다. 그 결과 표 2에서와 같이 D1/D3, S2/S1는 전문의의 평가 점수와 상관관계가 없는 것으로 나타났다( $p = 0.073, 0.599$ ). 나머지 8개의 파라미터는 통계적으로 유의( $p < 0.01$ )한 상관 계수(-0.36부터 0.60까지)를 보였다.

표 2에서 D1/D3, S2/S1은 상관관계가 없는 것으로 나타났는데 그 이유는 다음과 같다. 비공 간 거리의 분석에서 D1, D2,

D3으로 형성되는 삼각형이 D1 = D3 인 이등변 삼각형을 이룰 때, 변형이 되었음에도 불구하고 1에 가까운 값으로 계산되므로 오류를 나타낼 수 있다. 비공 면적의 분석에서도 S2/S1은 대칭성 및 비공의 모양과는 관계없는 면적의 비이므로 1에 근접하다고 정상으로 볼 수 없다. 이를 토대로 D1/D3, S2/S1을 제외한 나머지 8개의 파라미터를 임상적의 평가를 위한 인자로 설정하였다.

표 2. 각 파라미터와 전문의 평가점수와와의 상관관계 및 유의성

Parameter	correlation coefficient	p-value
$ \theta 1 - \theta 2 $	-0.36	0.0008
$\theta 1/\theta 2$	-0.39	0.0002
D2/D1	-0.45	<.000
D2/D3	-0.46	<.000
D1/D3	0.19	0.0725
S2/S1	0.06	0.5985
S3/S1	0.46	<.000
S3/S2	0.45	<.000
S3/(S1+S2)	0.60	<.000
S3/(S1+S2-S3)	0.57	<.000

신경망 모델의 경우, 입력노드는 구순열 코변형의 평가 영향을 미치는 인자들이고, 출력노드는 임상적의 평가 점수와 같이 0에서 100점 사이의 값이 되도록 하였다. 은닉층의 노드 수에 따라 결과에 영향을 미칠 수 있으므로 노드 수를 2에서 9까지 변형해 가며 학습을 수행하였다. 학습 데이터로는 전문의 2명으로부터 얻은 총 172개의 모든 데이터를 사용하였다.

신경회로망의 분석 결과, 일반인 간의 신경망 점수 차이는 평균이  $3.1 \pm 2.2$ 였다. 이것은 전문의 2인의 평균 평가 점수의 차이가  $9.52 \pm 11.50$ 인 것과 비교해 볼 때, 전문의들의 주관적인 평가보다 차이가 약 3배 이상 줄어들었음을 보여준다.

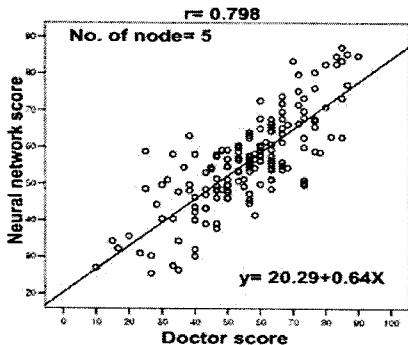


그림 2. 전문의와 은닉층의 노드 수가 5인 신경망 모델의 평가 점수 간의 상관관계

끝으로 전문의 평가 점수와 신경망을 이용한 점수 사이의 상관관계를 비교해 보았다. 전문의의 평가 점수와 신경망 추정치를 이용한 것의 상관관계를 살펴본 결과, 각각의 노드 수가 2에서 9로 증가함에 따라 상관 계수는 각각 0.701, 0.710, 0.758, 0.798, 0.781, 0.792, 0.834, 0.808로 나타났다. 노드 수가 8일 경우 가장 높은 상관 계수를 보였다. 그러나 히든 노드 수가

입력 노드 수와 같거나 더 클 경우 overfitting 되는 경우가 발생할 수 있다[5]. 따라서 은닉층의 최적의 노드 수는 5로 결정하였다. 그림 2는 은닉층의 노드 수가 5일 때의 상관관계를 보여주는 그래프이다. 그림 2로부터 신경망을 이용한 상관 계수 추정치로 전문의의 평가에 근접한 결과를 얻을 수 있었다.

#### 4. 결론

구순열로 인한 코 변형에 대한 평가는 주로 전문의에 의해 주관적인 평가로 판단되어 왔기 때문에 수술 결과의 만족도와 새로운 성형 수술 방법에 대한 효과나 만족도의 정량적 평가가 쉽지 않았다. 이는 본 연구에서도 3명의 전문의 중 1명은 다른 전문의 2명과 차이가 나타난 것으로 확인할 수 있었다. 또한 3명 전문의의 3회 평균점수에 대한 재현성도 컴퓨터를 이용한 방법의 재현성에 비해 낮았다.

본 연구에서는 이를 간단한 방법으로 객관적이고, 정량화하기 위하여 환자의 사진 영상만을 이용하여 비공 장축의 각도, 비공 간의 거리, 그리고 면적을 통한 코 변형의 비대칭성을 분석하여 비전문의도 쉽게 구순열 코변형 정도를 판단할 수 있도록 정량화된 값으로 표현할 수 있는 방법을 제안하였다. 또한 신경회로망을 이용한 컴퓨터 방법과 전문의의 평가점수가 높은 상관관계 ( $0.798, p < 0.000$ )를 가지는 것을 확인할 수 있었다. 본 연구 결과를 토대로 구순열 코변형 환자의 사진을 통하여 코변형 정도를 비공 중점 간의 거리와 비공간의 면적비 및 각도 등으로 코변형 정도를 정량화 할 수 있었으며, 코변형의 정도를 손쉽게 평가할 수 있는 지표를 마련할 수 있을 것으로 기대된다.

본 연구는 2004년 산업자원부 산업기술개발사업으로 수행되었음. (과제:10017058)

#### 참고 문헌

- [1] Anastassov Y, Chipkov C. "Analysis of nasal and labial deformities in cleft lip, alveolus and palate patients by a new rating scale", preliminary report. J Craniomaxillofac Surg. vol. 31, pp. 299-303, Oct, 2003.
- [2] Lo LJ, Wong FH, Mardini S, Chen YR, Noordhoff MS. Assessment of bilateral cleft lip nose deformity: a comparison of result as judged by cleft surgeons and laypersons. Plast Reconstr Surg. vol. 110, pp. 733-738, Sept, 2002.
- [3] Asher- McDade C, Roberts C, Shaw WC, Gallager C. Development of a method for rating nasolabial appearance patients with clefts of the lip and palate. Cleft Palate Craniofac J. vol. 28, pp. 385-390, Oct, 1991.
- [4] Coghlan BA, Laitung JK, Pigott RW. A computer-aided method of measuring nasal symmetry in the cleft lip nose. British Journal of Plastic Surgery. vol. 46, pp. 13-17, Jan, 1993.
- [5] Lawrence S, Giles CL, Tsoi AC, Lessons in Neural Network Training: Overfitting May be Harder than Expected, Proceedings of the Fourteenth National Conference on Artificial Intelligence, AAAI-97, AAAI Press, Menlo Park, California, pp. 540- 545, 1997.