

## 얼굴기미검출을 위한 중첩신경망의 전이학습에서의 한가지 망구성방법

심천룡, 리명철

경애하는 최고령도자 김정은동지께서는 다음과 같이 말씀하시였다.

《첨단과학기술분야에서 세계적경쟁력을 가진 기술들을 개발하기 위한 투쟁을 힘있게 벌려야 합니다.》

우리는 얼굴기미를 검출하는데 중첩신경망의 전이학습을 리용하여 중첩신경망의 한가지 형태인 GoogLeNet[3]의 마지막단계에서 얼굴의 기미와 정상피부를 식별하기 위한 층들을 새로 도입하는 방식으로 망을 구성하고 실험을 통하여 정확성을 검증하였다.

선행연구[1]에서는 기미검출을 위하여 GLCM화상으로부터 계산한 피부질평가용통계량들을 특징량으로 사용하였으며 기미가 있는 영역과 없는 영역을 판정하는데 지지벡터르기계(SVM)분류기를 리용하였다. 그러나 이 방법으로는 머리칼이 있는 영역까지 포함하여 얼굴화상의 모든 영역에서 기미점을 검출하는 경우 10%이상의 오류률을 얻었다.

선행연구[2]에서는 여러가지 음식물화상을 식별하는데 중첩심층신경망을 리용하였으며 음식물화상분류에 심층신경망이 효과적으로 리용될수 있다는것을 보여주었다.

본문에서는 GoogLeNet에 대한 전이학습을 리용하여 학습시킨 중첩신경망에 의한 얼굴화상의 기미영역판정에 대하여 논의하였다.

GoogLeNet연산그래프에서 최상단의 전결합층을 제거하고 그우에 4층으로 된 전결합층을 쌓아 분류기를 구성하였으며 측면조명을 반영한 비선형밝기변화방식에 의한 훈련자료증식, 목적하는 종류의 오류률감소에 무게를 반영한 손실함수에 대하여 연구하였다.

실험결과는 선행연구[1]와 같이 화상으로부터 특징량을 계산하고 그 특징량을 리용하여 SVM분류기로 분류하는것에 비하여 화상을 직접 입력하고 중첩신경망을 리용하여 분류할 때의 성능이 훨씬 높다는것을 보여준다.

8Mpixel화소점 이상의 얼굴사진 약 1만장으로부터 선행연구[1]에서 지정한 방법으로 기미가 있는 영역과 없는 영역의  $299 \times 299$ pixel화상을 각각 25 000장, 120 000장씩 얻은 후 분류기가 출력하는 믿음도값에 기초하여 믿음도가 낮은 화상에 대하여 수동으로 확인하는 방법으로 기미가 있는 화상과 기미가 없는 화상으로 분류하였다.

다음 비선형국부밝기변화와 히스토그램에 기초한 대조도변화를 통하여 훈련자료의 개수를 증가시켰다.

다음의 비선형국부밝기변화식은 각이한 측면조명을 모형화한것이다.

$$f(i, j) = (i, j) \cdot (\cos \alpha, \sin \alpha) = i \cos \alpha + j \sin \alpha \quad (i = 1, \dots, m; j = 1, \dots, n)$$

$$f_{\min} = \min_{i, j} f(i, j), \quad f_{\max} = \max_{i, j} f(i, j)$$

$$b_{i, j} = b_{\max} - (b_{\max} - b_{\min}) \cdot [2(f(i, j) - f_{\min}) / (f_{\max} - f_{\min}) - 1]^2$$

$$R'_{i, j} = \min\{R_{i, j} \cdot b_{i, j}, 255\}, \quad G'_{i, j} = \min\{G_{i, j} \cdot b_{i, j}, 255\}, \quad B'_{i, j} = \min\{B_{i, j} \cdot b_{i, j}, 255\}$$

여기서  $\alpha$ 는 밝기조절을 위한 방향각으로서  $[0, 2 \times \pi]$ 사이의 값이다.  $b_{\min}$ ,  $b_{\max}$ 는 각각 최소, 최대밝기승수이고  $R_{i, j}$ ,  $G_{i, j}$ ,  $B_{i, j}$ 는 각각 원본화상의  $(i, j)$  화소점에서의 RGB색값

이며  $R'_{i,j}$ ,  $G'_{i,j}$ ,  $B'_{i,j}$ 는 밝기변환을 진행한 후의 색값이다.

대조도변환공식은 다음과 같다.

$$i_{\text{new}} = \left\lfloor \frac{I_{\text{max}} - I_{\text{min}}}{M - L}(i - L) \right\rfloor + I_{\text{min}} \quad (L \leq i \leq M), \quad i_{\text{new}} = I_{\text{min}} \quad (i < L), \quad i_{\text{new}} = I_{\text{max}} \quad (i > M)$$

여기서  $i$ 와  $i_{\text{new}}$ 는 각각 대조도변환을 하기 전과 진행한 후의 색수준값이다.  $I_{\text{min}}$ ,  $I_{\text{max}}$ 는 변환후의 최소, 최대값이고  $L$ ,  $M$ 은 변환해야 할 수준범위의 최소, 최대값이다.

다음으로 망의 구조변경에 대하여 보자.(그림)

일반적으로 화상분류를 위한 중첩 신경망들은 여러개의 중첩층을 쌓고 중첩층의 최종결과를 전역선택연산을 거쳐 1차원벡토르형식으로 변환하며 분류하려는 클래스개수와 같은 출력을 가지는 전결합층을 쌓고 전결합층의 결과에 대한 유연최대함수를 계산하여 매 클래스에 속하는 확률을 얻는다.

우리는 GoogLeNet의 계산그래프에서 전역선택연산보다 우에 있는 층들을 없애고 2개의 클래스분류를 위한 5층으로 된 전결합층을 쌓았다.

학습시에는 전역선택연산보다 아래에 있는 층들을 동결시키고 전역선택연산보다 우에 있는 층들의 파라미터들만 학습시켰다.

기미검출에서 오유는 기미가 있는 영역을 기미가 없는 영역으로 판단하는 1종의 오유와 기미가 있는 영역을 없는 영역으로 판단하는 2종의 오유로 갈라볼수 있다. 현실적응용에서는 우의 두가지 오유중에서 보다 중요성이 높은 오유가 정해질수 있다.

우리는 1종의 오유 즉 기미가 있는 영역을 기미가 없는 영역으로 판단하는 오유를 줄이는데 무게를 두고 손실함수를 만들었다.

손실함수값  $J(\Theta) = -\frac{1}{m} \sum_{i=1}^m [\alpha \cdot y_s^{(i)} \log(\hat{p}_s^{(i)}) + y_n^{(i)} \log(\hat{p}_n^{(i)})]$ 은 교차엔트로피를 나타내는 량

이다. 여기서  $m$ 은 검증표본개수,  $y_s^{(i)}$ ,  $y_n^{(i)}$ 는 각각  $i$ 번째 검증표본이 기미있거나 없는 영역일 목적확률값이고  $\hat{p}_s^{(i)}$ ,  $\hat{p}_n^{(i)}$ 는  $i$ 번째 검증표본으로부터 계산한 기미가 있거나 없는 영역일 추정확률값이며  $\alpha$ 는 무게결수로서 1보다 크게 설정한다.

표 1. 학습파라미터탐색을 위한 탐색범위

학습파라미터	탐색공간
뭉음의 크기	[10, 50, 100, 500]
학습률	[0.01, 0.02, 0.05, 0.10]
활성화함수	[ReLU, ELU, Leaky_ReLU]
최적화방법	[Adam최적화, RMSProp최적화]
숨은층뉴트개수	N1[30, 50, 70], N2[10, 20, 30]

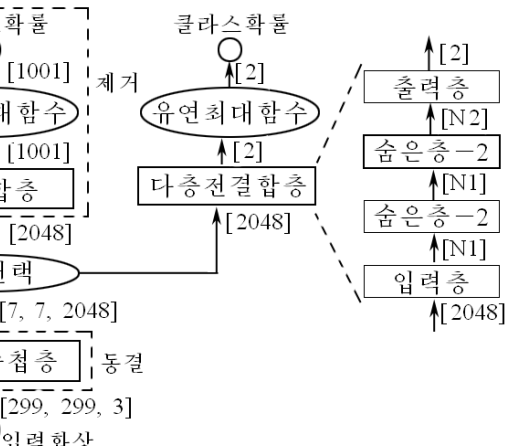


그림. 망구조변경

손실함수값을 최소화하는 무게파라미터와 편위파라미터는 소규모일괄학습에 의한 그라디언트하강법으로 구하였으며 훈련자료, 검증자료, 평가자료는 7:1:2의 비율로 분할하였다. 최적의 학습결과를 얻기 위한 학습파라미터탐색은 표 1과 같은 경우들에 대하여 진행하였다.

표 1에서의 활성화함수들은 다음과 같다.

$$h_{(w, b)}(X) = \max(Xw + b, 0) \quad (\text{ReLU})$$

$$\text{ELU}_\alpha(z) = \begin{cases} \alpha(\exp(z)-1), & z < 0 \\ z, & z \geq 0 \end{cases} \quad (\text{ELU})$$

$$\text{LeakyReLU}_\alpha(z) = \max(\alpha z, z) \quad (\text{Leaky\_ReLU})$$

먼저 최적화방법과 숨은층유니트개수를 고정한 상태에서 묶음의 크기와 학습률, 활성화함수의 최적값을 찾고 다음 찾은 값들을 고정한 상태에서 최적화방법과 숨은층유니트개수의 최적값을 찾았다.

학습도중에 30회이상 검증자료에 의한 손실함수값에서 개선이 없으면 학습을 중지시켰다.

평가자료에 의한 1종의 오류율과 2종의 오류율은  $E_1 = \frac{N_{FR}}{N_S}$ ,  $E_2 = \frac{N_{FA}}{N_N}$  와 같이 계산한다. 여기서  $N_{FR}$  는 기미가 있는 영역을 기미가 없는 영역으로 판단한 회수,  $N_S$  는 평가표본중에서 기미가 있는 영역의 화상표본개수,  $N_{FA}$  는 기미가 없는 영역을 기미가 있는 영역으로 판단한 회수,  $N_N$  은 평가표본중에서 기미가 없는 영역의 화상표본개수이다.

실험결과는 표 2와 같다.

실험결과는 선행한 방법에 비하여 전체적인 오류율을 감소시키면서도 목적하는 종류의 오류율감소에 무게를 주어 분류기를 학습시킬수 있다는것을 보여준다.

표 2. 실험결과

방법	$E_1$	$E_2$
SVM에 기초한 방법[1]	0.15	0.11
중첩신경망에 기초한 방법	0.02	0.05

## 참 고 문 헌

- [1] Chuan-Yu Chang et al.; Dermatological Sciences and Applications, 3, 1, 28, 2013.
- [2] P. Mc. Allister et al.; Computers in Biology and Medicine, 95, 2, 217, 2018.
- [3] M. M. Ghazi et al.; Neurocomputing, 235, 1, 228, 2017.

주체108(2019)년 9월 15일 원고접수

## A Network Construction Method in Transfer Learning of Convolutional Neural Network for Facial Spot Detection

Sim Chon Ryong, Ri Myong Chol

We propose a new network construction method to improve the accuracy of facial spot detection using transfer learning of convolutional neural network and verify the accuracy through experiments.

The results show that our new method is reasonable.

Keywords: transfer learning, spot detection, convolutional neural network