

딥러닝 기반의 이미지 분류를 이용한 패션 이미지 검색 웹사이트

Fashion Image Searching Website based on Deep Learning Image Classification

이학재*, 이석준, 최문혁, 김소영, 문일영

한국기술교육대학교 컴퓨터공학과

Hak-Jae Lee*, Seok-Jun Lee, Moon-Hyuk Choi, So-Yeong Kim, Il-Young Moon

Computer Science, Korea University of Technology and Education, Cheonan 31253, Korea

[요약]

기존에 존재하는 패션 웹 사이트에서는 상의, 하의 등의 품목에서는 한 가지 종류의 옷에 대한 검색결과만 보여주기 때문에 사용자가 원하는 옷에 대한 조합을 찾을 수 없다. 또 패션 시장이 성장함에 따라 소비자들은 다양한 패션 정보를 찾을 수 플랫폼을 요구하고 있다. 이러한 문제를 해결하고자 하여 딥러닝을 통한 이미지분류를 웹 사이트와 연동하고 SNS 기능을 접목하는 아이디어를 고안해냈다. 웹 사이트에 사용자가 본인의 이미지를 업로드하여 딥러닝 서버를 통해서 이미지의 특징을 파악하고 분류하여 저장한다. 사용자들은 저장된 정보를 가지고 여러 조합을 통해 원하는 이미지들을 검색할 수 있다. 또 SNS 기능을 통해 사용자간의 커뮤니케이션이 활발하게 이루어질 수 있다. 이를 통해서 기존에 존재하는 패션 관련 사이트의 문제를 해결하는 방안을 마련하였다.

[Abstract]

Existing fashion web sites show only the search results for one type of clothes in items such as tops and bottoms. As the fashion market grows, consumers are demanding a platform to find a variety of fashion information. To solve this problem, we devised the idea of linking image classification through deep learning with a website and integrating SNS functions. User uploads their own image to the web site and uses the deep learning server to identify, classify and store the image's characteristics. Users can use the stored information to search for the images in various combinations. In addition, communication between users can be actively performed through the SNS function. Through this, the plan to solve the problem of existing fashion-related sites was prepared.

Key Words: Computer vision, CNN, Deep learning, Image classification, Web

<http://dx.doi.org/10.14702/JPPE.2019.175>



This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Received 8 November 2019; Revised 12 November 2019

Accepted 13 November 2019

*Corresponding Author

E-mail: iiefund@koreatech.ac.kr

I. 서론

지난 10년간 국내 패션 시장 규모는 9조원 가량 성장해왔다. 10여년 동안 사람들이 패션에 대한 관심이 점점 증가함으로써 자신에게 어울리는 착장을 찾거나, 원하는 옷을 찾기 위해 패션과 관련된 웹 사이트나 어플리케이션 등의 다양한 매개체로 정보를 찾는 경우가 많아졌다. 또한 이러한 패션은 20대부터 40대 까지 다양한 연령층에서 관심이 쏠리고 있다. 그들은 SNS를 통해 정보를 찾고 공유하거나, 웹 사이트의 정보를 통해서 원하는 것을 찾는다.

사람들은 이러한 패션에 대한 정보를 다양한 매개체를 통해 찾음에도 불구하고, 시간이 지나면 지날수록 의류 트렌드가 급격하게 바뀌고, 어떤 조합으로 의류를 착용해야할지 모르는 사람들이 대다수 있어서 정보가 많음에도 불구하고 의류를 조합하고 착용하는데 어려움이 있다. 요즘은 사람들은 패션관련 웹 사이트에서 단일 품목으로 의류를 검색하며 정보를 찾고 있다. 또한 온라인 스토어나 다양한 패션 커뮤니티 사이트에서는 모델을 통하여 코디를 제시하지만, 대부분은 한 가지 결과만 보여주기 때문에 사용자가 여러 개의 조합을 찾기 힘든 단점이 있다. 대부분의 패션 관련 웹 사이트에서는 상의, 하의 등의 카테고리에서 존재하는 한 가지 종류나 색깔의 옷에 대한 검색결과만 보여주기 때문에 사용자가 상하의 별로 원하는 옷에 대한 조합을 찾을 수 없다.

이러한 문제를 해결하기 위해서 딥러닝의 CNN 기술을 통한 이미지 분류와 커뮤니티 기능을 활용하여 해결하는 방안에 대해서 고안하였다. 인터넷이 발달하고, 다양한 서비스가 생김에 따라 커뮤니티나 SNS등의 사람들과 소통이 가능한 매개체는 급격하게 늘어났고, 정보 공유에 있어서 좋은 방안이 되었다. SNS와 커뮤니티를 통해서 사람들은 다양한 주제를 서로 공유하는데, 이러한 SNS와 커뮤니티가 특정 주제에 포커스가 되어있으면 관련된 데이터를 수집하기가 비교적 편리하다. 이러한 특성을 이용하여 패션과 관련된 웹 사이트를 구현하여 다양한 정보를 사람들이 공유할 수 있도록 하였고, 또한 사람들이 다양한 조합을 통해 검색을 할 수 있도록 사람들의 착장 이미지를 수집하여 딥러닝 인공지능망 CNN을 통해 이미지들을 학습시키고 해당 이미지를 각각의 카테고리로 분류를 하고 데이터베이스에 데이터화 하여 저장할 수 있도록 하여 사람들이 쉽게 다양한 패션 조합을 검색할 수 있는 방안을 고안해냈다.

이에, 본 논문에서는 기존의 패션관련 웹 사이트의 문제점을 개선한 다양한 기능을 추가한 새로운 웹 사이트를 제안하였다.

본 논문의 구성은 다음과 같다. 제 II장에서는 웹서비스의

구성 기술에 대하여 알아본다. 제 III장에서는 딥러닝 인공지능망인 CNN을 활용한 이미지 분류 기술에 대하여 알아본다. 마지막으로 제 IV장에서는 본 논문의 결론을 맺는다.

II. 웹서비스 기술

구성한 웹 사이트는 프론트엔드와 백엔드로 구성되어있다. 프론트엔드에는 Node.js 기반으로 한 템플릿 엔진인 pug를 이용하여 웹 페이지 프론트엔드를 구현하고 웹 디자인을 적용하였으며, 백엔드에는 웹 서버로는 Nodejs를 이용하여 서버를 구현하였고, 데이터베이스는 Mongo DB를 이용하여 데이터베이스를 구현하였다.

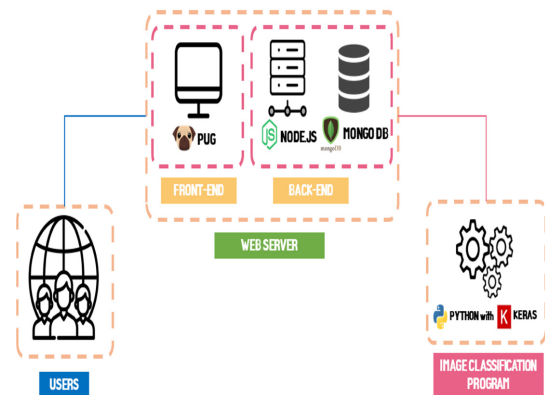


그림 1. 서버 구조

Fig. 1. Server structure.

구성한 웹 사이트는 그림 1과 같은 서버 구조를 가진다. Node.js에서 사용하는 템플릿 엔진은 pug, ejs, handlebars 등 많이 존재한다. 여기서 pug를 사용하는 이유로는 html문법보다 코드가 간소화되었으며, 컴파일한 후 html문서를 렌더링하는 형식이라 생산성이 높아지기 때문에 사용하였다.

페이지 구성으로는 회원가입, 회원탈퇴, 검색, 통계, 자유게시판, 공지사항, 사진 업로드, 마이페이지, 사진 상세 페이지로 구현하였다.

그림 2에서 보면 메인 화면에서는 사용자가 사진 업로드를 통해 올린 사진들을 보여주도록 하였고, 각각의 사진마다 추천 기능을 넣어서 가장 인기가 많은 사진을 보여주는 추천순과 업로드 시간을 기반으로 한 최신순으로 사진을 정렬할 수 있도록 구현하였다. 또한 사진을 클릭해서 사진을 올린 작성자와 sns 아이디, 의류 사진, 댓글 기능을 통해 sns와 같은 기능

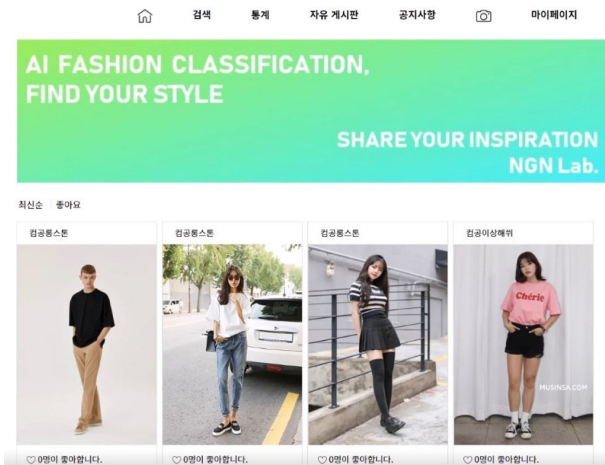


그림 2. 메인 화면

Fig. 2. Main page.

을 추가하였다. 사진 상세보기 기능을 통해서 사람들이 해당 게시글을 좋아하면 해당 글을 추천하고 서로 감상평이나 피드백을 남겨 다른 사용자와 커뮤니케이션을 할 수 있다.

그리고 검색은 데이터베이스에 CNN을 통해 분류하여 저장된 이미지 데이터를 기반으로 사용자가 원하는 조합을 검색해서 코드를 확인할 수 있도록 구성하였다. 검색의 종류로는 7가지가 존재하며, 성별, 상의 종류, 상의 패턴, 상의 색상, 하의 종류, 하의 패턴, 하의 색상으로 구성되어있다. 상의의 종류는 긴팔 티셔츠, 반팔 티셔츠, 셔츠, 블레이저, 청자켓의 5가지로 분류되고, 하의의 종류로는 청바지, 면바지, 반바지, 숏치마, 롱치마의 5가지로 분류된다. 그리고 상하의별로 민무늬, 체크무늬, 줄무늬의 3가지로 분류된다. 그리고 색상은 총 13가지가 존재한다. 사용자들은 주어진 조건을 통해서 원하는 코드를 검색할 수 있도록 하였고, 데이터베이스 내부에서 저장된 검색 조건에 맞는 내용을 출력하도록 하였다. 해당 기능은 기존에 존재하는 단일 품목만 보여주는 웹 사이트보다 다양한 조합으로 검색할 수 있어서 핵심 기능으로 구현하였다.

통계 기능은 사용자가 올린 서버에 저장된 이미지들을 바탕으로 알고리즘을 이용해 어떤 의류가 웹 사이트 이용자에게 가장 인기가 많은지 각각의 종류에 대하여 순위별로 나타내었고, 또한 각각의 의류별로 어떤 의류와 매칭이 많이 되는지 순위별로 각각의 퍼센트를 나타내었다. 예를 들어, 반팔을 입은 남자는 어떤 하의와 가장 코드를 매칭을 많이 하는지 하의를 순위별로 확인할 수 있도록 하였다. 이러한 통계를 이용하여 어떤 조합이 가장 인기가 많은지 확인할 수 있도록 하였다. 그리고 자유게시판, 공지사항은 일반적인 커

뮤니티 사이트의 글쓰기, 수정, 삭제, 열람, 댓글의 기능을 구현하였고, 회원가입과 회원탈퇴를 통해서 회원정보를 DB에 저장하거나, 삭제하는 페이지를 구성하였다.

III. CNN기반의 이미지 분류 기술

이미지 분류의 경우에는 인공신경망의 CNN을 통해서 구현하였다. Fully Connected Layer 만으로 구성된 인공 신경망의 입력 데이터는 1차원 형태로 한정된다. 사진 데이터로 전 연결 신경망을 학습시켜야 할 경우, 3차원 데이터를 1차원으로 평면화를 진행하여야 한다. 사진 데이터를 평면화 시키는 과정에서 공간 정보가 손실될 수가 있다. 이미지의 공간 정보를 유지한 상태로 학습이 가능한 모델인 CNN을 사용하였다.

A. CNN 학습 수행 과정

텐서플로의 딥러닝 모델 설계와 훈련을 위한 API인 keras를 이용하여 CNN 모델을 생성하고 학습 데이터를 입력하여 반복 학습을 진행하였다. 학습 데이터로는 얼굴을 통해 성별을 인식할 수 있고, 5종류의 상의와, 5종류의 하의, 3종류의 무늬에 관한 사진들로 구성되어있다. 일단 좋은 학습 결과를 얻기 위해서 몇 가지 작업을 진행하였다. 먼저, 추측할 데이터 모델과 학습한 데이터 모델이 최대한 유사한 형태를 보이도록 구성하였다. CNN 모델에 맞는 이미지를 처리하기 위해 전처리 과정을 거쳤는데, 여기서는 이미지의 scale 범위를 동일하게 만들고 사용될 이미지의 크기와 채널을 동일하게 구성하였다. 이미지가 유사할 경우에는 높은 정확도를 가지는 것으로 확인되었다. 학습 데이터로는 인터넷 상에 존재하는 전신 사진을 통해서 각각을 얼굴, 상의, 하의별로 나누어서 잘라내어 얼굴, 상의, 하의 모델에 적용시킬 수 있도록 이미지를 수집하였다. 실제로 업로드되는 이미지들 또한 이와 같은 형태일 것으로 예측하여 정확도가 높아야 이미지 분류에 의미있는 결과를 가져올 것이고, 검색과 통계에 좋은 결과를 가져오는 효과를 기대할 수 있기 때문에 이와 같은 과정을 거쳤다. 패션 이미지는 사람들이 서있는 자세와 모양에 따라 옷의 형태나 모양이 달라질 수 있고, 이미지에서 보이는 옷의 특징이 다르게 보일 수 있기 때문에 실제 이미지에서 사람들이 주로 취하는 자세와 모양을 가지는 이미지를 활용하였다.

CNN 모델은 합성곱 필터가 존재하는 합성곱 계층과 풀링을 진행하는 계층으로 반복하여 구성된다. 먼저 합성곱 계층

에서는 입력 데이터를 필터를 통해서 합성곱 연산을 수행하였다. 합성곱 처리 결과로부터 Feature Map을 만들었다. 필터링이 적용된 이미지는 새로운 형태의 이미지로 변환되고 이를 다음 계층인 풀링 레이어로 전달한다. 풀링 레이어는 컨볼루션 레이어의 출력 데이터를 입력으로 받아서 Feature Map의 크기를 줄이거나 특정 데이터를 강조하는 용도로 사용된다. 풀링 레이어를 처리하는 방법으로는 Max Pooling과 Average Pooling, Min Pooling이 있는데, 여기서는 Max Pooling을 통해 국소적인 부분에서 가장 큰 값을 취함으로써 이미지를 변화시켰다. 그리고 중간에 Dropout을 통해 계층을 통과한 값들의 일부를 사용하지 않고 지나감으로 학습 데이터에 과적합되는 현상을 방지하였다. 이렇게 Convolution layer와 풀링 레이어의 Max Pooling을 반복적으로 stack을 쌓아서 마지막 출력 층에 도달한 값을 총 5가지의 클래스로 분류하였다. 이 과정을 통해서 학습 데이터들의 특징을 학습하고 새로운 이미지에 대해 클래스를 추측하는 이미지 분류가 완성되었다.

B. 이미지 분류 알고리즘

모델에 사용된 훈련셋은 각 클래스 당 약 1000여장으로 훈련셋이 시험셋의 특성을 충분히 반영하기에는 부족한 수치이다. 그러나 이러한 한계점을 보완하기 위해 컨볼루션 신경망의 성능을 향상시키기 위해 데이터 부풀리기 기능을 활용하였다. 데이터 부풀리기는 ImageDataGenerator 라는 함수를 이용하였으며, 데이터 부풀리기를 통해서 이미지의 회전, 이동, 확대, 축소, 뒤집기 등 다양한 방법을 통해서 새로운 학습 데이터들을 생성하여 학습을 진행하였다. 학습 결과를 통해 정확도가 90% 이상 나오는 모델을 사용하였다. 이미지 분류 과정은 전신의 사람 인식부터 시작하였다. 하나의 신경망을 전체 이미지에 적용하기 위해서 YOLO라는 알고리즘을 활용하였다. 이 신경망은 이미지를 영역으로 분할하고 각 영역의 Bounding Box와 확률을 예측하고, 이런 Bounding Box는 예측된 확률에 의해 가중치가 적용된다. 앞서 훈련된 모델을 이용하여 사람을 인식할 수 있도록 응용하였다. 인식된 사람은 Bounding Box로 표현되며 Bounding Box의 좌표를 분리하여 사람을 추출하였다. 추출된 사람에게서 필요한 정보인 얼굴(성별), 상체, 하체의 3가지 부분을 아래와 같이 그림 3, 그림 4, 그림 5와 같은 모습으로 가져왔다.

얼굴 인식의 경우에는 Haar Cascade라는 알고리즘을 사용하였다. Haar Cascade는 머신 러닝기반의 오브젝트 검출 알고리즘이다. 해당 알고리즘은 비디오 또는 이미지에서 오브젝트를 검출하기 위하여 사용된다. 직사각형 영역으로 구성

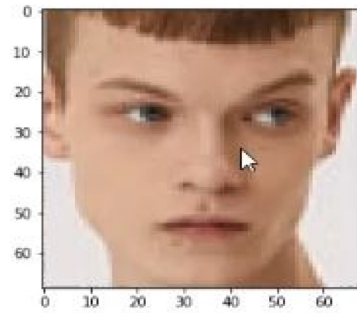


그림 3. 얼굴 인식

Fig. 3. Face recognition.



그림 4. 상체 인식

Fig. 4. Top recognition.



그림 5. 하의 인식

Fig. 5. Bottom recognition.

되는 특징을 사용하기 때문에 픽셀을 직접 사용할 때 보다 동작 속도가 빠르다. 밝은 영역의 픽셀 값의 합과 어두운 영역의 픽셀 값의 합의 차이를 통해서 이미지의 특징을 찾는다. 얼굴 인식에 사용되는 수치를 활용하여 얼굴에 Bounding Box로 표현하며, 해당 Bounding Box의 좌표를 추출하여 성별 분류 모델에 적용하는 방식을 사용한다. 그리고 상체와

하체의 경우에는 인체의 비율을 통해 사진을 분리하여 사용한다. 인체 비율의 경우에는 여러 개의 경우를 통해서 얼굴을 제외한 사람의 전신 이미지를 상체 하체로 나누고 적절하게 분리된 수치를 찾아 적용하였다. 추출된 사람의 전신이미지에서 이미지의 높이를 머리 상체 하체 별로 1, 2.4, 3.3의 비율로 분리하였다. 사람마다 비율이 다르기 때문에 정확한 이미지를 얻는 것은 불가능하다. 하지만 이미지 속에 있는 옷의 종류와 특징, 무늬가 있으면 이미지를 분류하는데 큰 문제는 없는 것으로 확인되었다. 색상의 경우에는 분리형 군집화 알고리즘인 K-means 알고리즘을 통해서 이미지 내에서 k개의 평균을 만들어 클러스팅 하는 방식을 응용하였다. RGB 값으로 이미지에 나타나는 색상을 K개로 분류하고, 가장 빈도가 많은 RGB값을 옷의 색상으로 지정하였다. 13개의 계통의 색상은 파이썬에서 제공하는 사전의 형태로 저장하였다. K-means 알고리즘으로 도출된 RGB값과 사전에 있는 RGB값을 비교하여 가장 근접한 색상의 계통으로 값을 지정하였다.

```

brown6
남자
반팔티 민무늬 black
면바지/슬랙스 민무늬 brown
{
  "sex": "남자",
  "top": "반팔티",
  "topPattern": "민무늬",
  "topColor": "black",
  "bottom": "면바지/슬랙스",
  "bottomPattern": "민무늬",
  "bottomColor": "brown",
  "fileUrl": "1(2).jpg"
}

```

그림 6. 그림 2에 있는 첫 번째 사진의 분류 결과

Fig. 6. Classification result of the first picture in Fig. 2.

그림 6은 그림 2에 있는 첫 번째 사진을 업로드 하여 수행한 결과로, 그림 3, 그림 4, 그림 5와 같이 얼굴, 상의, 하의로 추출되어서 그림 6과 같은 결과가 도출되었다. 그림 6의 결과는 그림 2에 있는 사진과 같이 남자, 반팔티, 민무늬, black, 면바지/슬랙스, 민무늬, brown과 같이 결과가 나오게 되었다.

이렇게 얼굴, 상의, 하의를 추출하고 CNN 모델을 통해 학습된 내용을 바탕으로 추측 결과를 JSON 파일 형식으로 저장하여 웹서버에 저장하였다.

웹 서버는 해당 JSON 파일의 내용과 사진을 그림 7과 같

```

{ "_id" : ObjectId("5d8a77566498465e70eae27f"), "fileUrl" : "upload\\image\\5d8a771fde4a7964683eeaca\\1(2).jpg", "sex" : "남자", "top" : "반팔티", "topPattern" : "민무늬", "topColor" : "black", "bottom" : "면바지/슬랙스", "bottomPattern" : "민무늬", "bottomColor" : "brown", "user" : [ ObjectId("5d8a771fde4a7964683eeaca") ], "like" : 0, "count" : 276, "comments" : [ ], "nickName" : "컴공콩스톤", "_v" : 0 }

```

그림 7. JSON 파일이 저장된 데이터베이스

Fig. 7. Database where JSON file is stored.

이 데이터베이스에 저장하였다.

IV. 결 론

해당 연구를 통해 기대하는 효과는 CNN을 통해 학습된 이미지를 통해 대량의 이미지 데이터 수집이다. 기존의 패션 관련 데이터 검색이나 수집은 특정 키워드나 단일 품목의 검색이 주를 이루지만, 본 연구와 같은 방식을 이용하여 이미지를 몇 가지의 종류로 분류하고 데이터화 하여 다양한 방법으로 검색할 수 있는 효과를 기대한다. 이러한 이미지 분류 방법은 패션 이미지 분류 뿐만 아니라 다양한 분야에서 활용할 수 있기 때문에 활용가치가 방대할 것이라고 기대한다. 해당 연구에 대해 보충할 부분도 어느 정도 존재하는데, 먼저 이미지 분류 모델의 완성도가 있다. 해당 모델들은 데이터셋이 부족한 상태로 학습이 진행되어서 아쉬운 점이 있다. 하지만 클래스당 학습 이미지 수를 수 십만장 단위로 학습할 경우에는 다양한 옷의 종류의 분류와 정확도를 기대할 수 있을 것이다. 그리고 해당 모델은 사람의 다양한 자세에 대해 검출할 수 없는데, 상체와 하체를 개별적인 객체로 인식하여 이미지 인식을 할 수 있으면 다양한 자세의 사람들의 이미지를 인식하고 분류할 수 있는 기능 구현이 가능할 것이라고 생각한다.

감사의 글

본 과제(결과물)는 교육부와 한국연구재단의 재원으로 지원을 받아 수행된 사회맞춤형 산학협력 선도대학(LINC+) 육성사업의 연구결과입니다.

참고문헌

- [1] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: unified, real-time object detection," in *Proceedings of the IEEE*, pp. 1-10, 2016.
- [2] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems 25*, 2012.
- [3] A. A. Akash, A. S. Mollah, and M. A. H. Akhand, "Improvement of haar feature based face detection in OpenCV incorporating human skin color characteristic," *Journal of Computer Science Applications and Information Technology*, vol.1, no. 1, pp. 1-8, 2016.
- [4] S. J. Kim, Coding Chef's 3 Minute Deep Learning Keras Flavor, Hanbit Media, 2018.
- [5] H. Y. Jo, Node.js textbook, Gilbut, 2018.



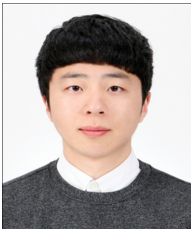
이 학 재 (Hak-Jae Lee)_학생회원

2014년 2월 ~ 현재 : 한국기술교육대학교 컴퓨터공학과 학사과정
<관심분야> 웹, 딥러닝, 컴퓨터 비전



이 석 준 (Seok-Jun Lee)_학생회원

2014년 2월 ~ 현재 : 한국기술교육대학교 컴퓨터공학과 학사과정
<관심분야> 딥러닝, 데이터마이닝



최 문 혁 (Moon-Hyuk Choi)_학생회원

2014년 2월 ~ 현재 : 한국기술교육대학교 컴퓨터공학과 학사과정
<관심분야> 웹



김 소 영 (So-Yeong Kim)_학생회원

2014년 2월 ~ 현재 : 한국기술교육대학교 컴퓨터공학과 학사과정
<관심분야> 웹



문 일 영 (Il-Young Moon)_종신회원

2000년 2월 : 한국항공대학교 항공통신정보공학과 졸업(공학사)
2002년 2월 : 한국항공대학교 대학원 항공통신정보공학부 졸업(공학석사)
2005년 2월 : 한국항공대학교 대학원 정보통신공학과 졸업(공학박사)
2004년 ~ 2005년 : 한국정보문화진흥원 선임연구원
2005년 3월 ~ 현재 : 한국기술교육대학교 컴퓨터공학과(교수)
<관심분야> 모바일 프로그래밍, 무선 인터넷, 모바일 IP