*Abstract*—Nowadays, as kinds of medicine become more varied, consumers require more detailed information of medicine such as name, effect, or any side effects of medicine. For example, some consumers may need to avoid getting certain kinds of medicine. Based on those demands, we made an application which provides information of medicines to customers by image searching, based on machine learning.

I. Introduction

In real life, information of medicine has become more important. First, for example, if some different kinds of medicines are mixed by chance, it is obviously needed to separate medicines by related kinds of medicines. Second, consumers should be able to choose medicine. Even if it might be risky to take certain kinds of medicine arbitrarily, it would be helpful in terms of letting consumers be able to have more positive choices.

To satisfying these purposes, we made an application to identify medicines by image. With rapid growth of machine learning method, now we are able to apply machine learning method in application. We used machine learning method for image comparison to identify medicines and provide the information of it. Using this application, customers would get not only information about medicines, but also more choices in their life.

Users can either take a picture of a pill themselves or retrieve a pill picture from a gallery to find out what the pill is. This search is useful when user does not have information such as the name or ingredients of a pill and user has pictures or the substance of the pill. Based on machine learning and pill image data, the app can identify which pill is in the entered image. To make app identify better, app highlight certain section of camera display in which medicine should be.

In given picture, ROI(Region of Interest) should be caught correctly, which means to know where the pills are in picture. Without additional function, we highlight center part of camera view so that users can focus on setting pills in the center of picture. This way not only gives more exact ROI, also let users make better quality of picture. In case of pictures loaded from gallery, same thing is applied. Users can move and zoom in and out to make pills be on center of the picture.

We use CNN method for pill recognition with computer vision. CNN, Convolution Neural Network is most dominant method in computer vision. Among the open source model of CNN, we take the GoogleNet Inception and do ‘Transfer Learning’ for this model. Because this model has guaranteed performance in many sections and have a lot of reference we can look for which means those kernel would be working in our subject, so we don’t make new convloution network and learn new kernels.

II. Related Work (e.g., existing studies)

Before we start, there were some applications about providing information about medicines. ‘Smart DUR’, ‘Searching medicines’, ‘Recommendation app by using prescription and barcode’, ‘KIMS Mobile’ are similar applications in terms of searching. These applications provide information about the ingredients of the medicine, the interaction with the food, and so on. And ‘Pill Scanner’ application also has same functions, but this application’s main function is scanning medicine by photos. Based on AI, this application analyzes the medicine(photo) and shows information about the medicine.

Smart DUR has these functions. First, users can search medicines they can get information about medicine, how to take drugs, and cautions about medicine. Second, applications tell whether the medicine has any problems. Lastly, if users do not know the medicine name, user can search the medicine by medicine’s shape, color, etc. And this application saves history that users had searched. Our application also needs this function.

Pill Scanner’s main function is searching medicine by image analyzing. This function is what we want to make mainly. Users can search the medicines by name. But this application has weakness about searching. When search by image, it doesn’t recognize the medicine well. This application is not enough in performance, but these functions will help us for making our application.

Google OCR can be used in Google Drive, without a any installation. Upload at the google drive with the file that is end with .jpg, .png, .pdf, etc. Open these files with google docs, then users can see the photos with letters that is in photos. We can use this method with our works

Google Inception-v3 model is open source CNN model from google. This is not the latest model, but has enough, so guaranted performance for many kinds of images that we choose this model as default. We do transfer learning with this trained model, so it would be a kind of fine tuning for our task, pill image recogintion. More details about this model are in the Methodology part.

There is an interesting paper on pill image recognition by Yu Wang : PILL RECOGNITION USING MINIMAL LABLED DATA. Even though we do not make our own algorithm for image recognition, this paper is still meaningful because it provdies how to deal with minimal datasets for machine learning, pill recognition. This paper makes some categories which is important for pill recognition in terms of computer vision, such that how pills’ colors can be clustered in machine learning using K-means and what kinds of pills’ shapes are regognized in computer vision. So we could learn how to deal with minimal pill datasets.

Reference

Smart DUR : http://smartdur.firstdis.co.kr/

Pill Scanner : https://play.google.com/store/apps/details?id=com.aitrics.pillscanner.clinical&hl=ko

Google OCR : https://cloud.google.com/vision/docs/ocr

Tensor Flow : https://www.tensorflow.org/?hl=ko

Google inception-v3 model : <https://arxiv.org/abs/1512.00567>

PILL RECOGNITION USING MINIMAL LABLED DATA : <https://www.computer.org/csdl/proceedings/bigmm/>2017/6549/00/07966771.pdf

III. Datasets

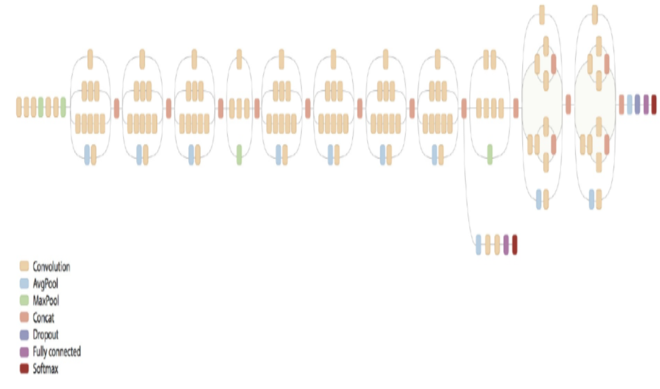
First we decided what kinds of pills would be used. It was the difference of color and shape that the most important thing for this choice of pills. It is to make it simple to distinguish pills. And the next is how popular that pill is. Because popular pill has two advantage : easyness to gaher images and usefullness of software in reality. So we selected two kinds of pilss, one is aspirin which is white and circle shaped and the other is advil, red color and oval shaped. Then we gathered 30 images for each pill to test the performance of transfer learned model. These are in the github repostitory below.

(https://github.com/kgr2140/Medetector/tree/ai\_mixroa/AI/tools/img/pills)

IV. Methodology

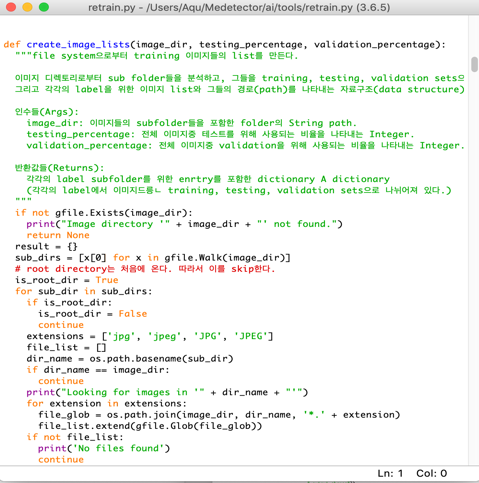
1. **Google inception v3**

This model is a kind of CNN model from google. CNN, Convolution Neural Network is a dominant deep learning method in computer vision and image recognition. A CNN consists of an input and an output layer, as well as multiple [hidden layers](https://en.wikipedia.org/wiki/Multilayer_perceptron#Layers). The hidden layers of a CNN typically consist of convolutional layers, pooling layers, fully connected layers and normalization layers Description of the process as a [convolution](https://en.wikipedia.org/wiki/Convolution) in neural networks is by convention. Mathematically it is a [cross-correlation](https://en.wikipedia.org/wiki/Cross-correlation) rather than a convolution. This only has significance for the indices in the matrix, and thus which weights are placed at which index.



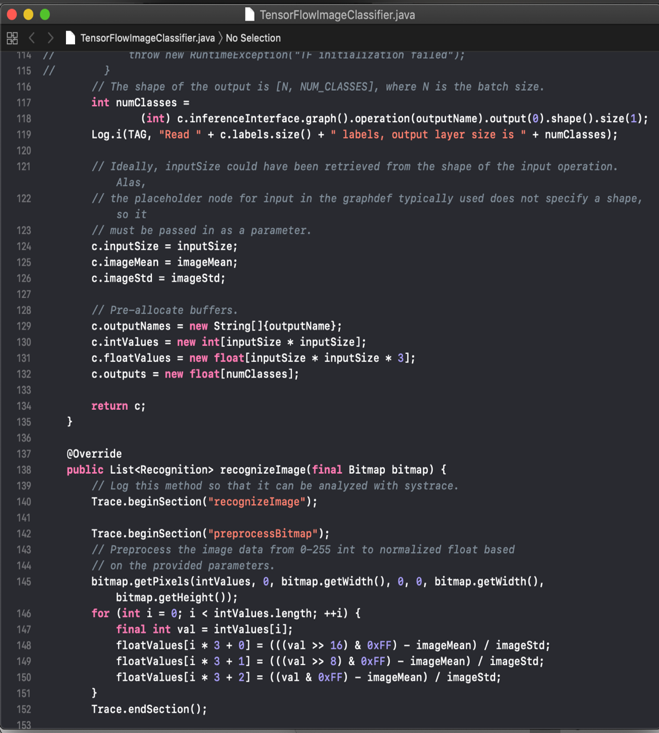
1. **Transfer learning**

Basically we use pre-tranined google inception v3 model, which means we just borrow previously leanred filter set to transfer learning for our own data set. This approach makes CNN model be able to learn specific pattern according to our own labels.



1. **Graph Stripping**  
   Now, to make the tranfer-leanred graph be able to work in android mobile application, additional transform process is needed. This process is provided by Tensorflow library, called graph stripping. It removes some nodes which is not used in mobile platform so that it can work in mobile, smartphone without extra server.
2. **Interact with android**

There is a java API in Tensorflow, Classifire.java file. It brings results which are obtained from the learned graphs as performing a interface role between user input and machine learning graphs in the overall structure.



V. Evaluation & Analysis

Evaluation : Genrally it works so good in our data set which has different colors. About no less than 90% was the classfying rate for each pill, aspirin and advil, but in some specail cases, which has some other objects in picture such as colored medicine bottle and some medicine mark, it was 60-70 percent, more than 65% though.

Analysis : Above all, the number of data set, 30 images for each label, was too little. It was not enough to extract specific features from images. Since we found that exceptional lower cases have some stuff which has different color, we analyze that color was the most dominant parameter in cross validaing for 30 images.

VI. Discussion & Conclusion

1. 트랜스퍼 러닝에 대하여

: CNN 개념의 등장 이래로 이미지 인식에 대한 딥러닝 알고리즘은 급격한 발전을 이루었다. 이는 곧 프로그래머들이 완전히 새로운 모델 및 layer를 설계하지 않고도, 이미 학습된 모델에 their own dataset을 사용해 transfer learning을 하는 것만으로도 유의미한 performance를 이끌어낼 수 있게 되었음을 의미한다.

1. 모바일 범용성에 대하여

Stripped graph는 mobile platform에서 잘 작동하는 모습을 보였다. Mobile을 위한 Tensorflow lite가 아직 지원되지 않음에도 안정적인 구동과 퍼포먼스를 보여준다는 사실은 스마트폰 스펙의 상향평준화와 맞물려 딥러닝이 모바일 단에서 충분히 작동될 수 있음을 의미한다.