An End-To-End Crowdsourcing Platform For Pathological Images

Submitted by:
Ashutosh Yadav (B16CS005)
Chetan Prakash Meena (B16CS006)

Under the supervision of
Dr. Sumit Kalra
Department of Computer Science and Engineering,
Indian Institute of Technology Jodhpur,
Rajasthan

Department of Computer Science and Engineering
Indian Institute of Technology Jodhpur,
Rajasthan

CERTIFICATE

This is certify that the work contained in this entitled "**An End-To-End Crowdsourcing Platform For Pathological Images**" is a bonafide work of **Ashutosh Yadav (B16CS005)** & **Chetan Prakash Meena (B16CS006)** carried out in **Department of Computer Science and Engineering Indian Institute of Technology Jodhpur** under my Supervision and it has not been submitted elsewhere for a degree.

Supervisor:

Dr. Sumit Kalara
Department of
Computer Science & Engineering
IIT Jodhpur, Rajasthan
April 2019

Acknowledgement

First we would like to thank our guide Dr. Sumit Kalara for guiding us for this project and giving us opportunity to working on Web development and Machine Learning. We would also like to thanks our friends who initially helped us.

By:

Ashutosh Yadav
Chetan Prakash Meena

Contents

1. Introduction.....

2. Problem Formulation.....

3. Tools and API used.....
    3.1. Vuejs and Vuetify
    3.2. Firebase
    3.3. Twitter API
    3.4. Python
    3.5. Jupiter
    3.6. Keras
    3.7. Tensorflow

4. Methodology.....
    4.1. Overview
    4.2. Detailed Overview
    4.3 UML Diagram

4. Future Plans.....

5. Result and Conclusion.....

6. References.....

1. Introduction

In today's era more and more people are moving towards **Deep Learning** as its outcomes are unbelievable even sometime far better than human's imagination. In medical industry some software are giving extraordinary results but in some specific domain because these software are generally trained using deep learning and it requires huge amount of data which is not available for some domain. We are here creating a labelled dataset for pathological images so that people can use this dataset in there project to build more amazing project.

The only thing we can do is create an unlabelled dataset for pathological images but we can create an interface through which we can labelled our dataset using crowdsourcing. So we created a website for this interface.

2. Problem Formulation

Broadly our problem is divided into three parts:

1. Downloading Images from Twitter
2. Creating dataset of Pathological Images
3. Labelling this dataset using crowdsourcing

3. Tools and API used

We have uses following tools:

3.1. Vuejs and Vuetify

For frontend purpose

3.2. Firebase: Database, Storage

For Database, Storage, Authentication, Back-end Functions, Hosting

3.3. Twitter API

For downloading images and tweets from twitter

3.4. Python

For DL model and downloading images from twitter

3.5. Jupiter

For Running python script in part

3.6. Keras

For deep learning

3.7. Tensorflow

Keras backend for training

4. Methodology

4.1. Overview:

In this project, we proposed a deep learning based method using Convolutional Neural Network for image classification. These images are downloaded from Twitter using python script. For managing these images we created another python script. Through this script we can easily maintain these images.

After creating our dataset of pathological images our task is to create an interface through which people can access our dataset and give there ideas about these images so that we can labelled our dataset. We create this interface in form of website, we build this website using Vue platform as it is easy to handled. In the website users can also discuss with other users about a particular images and can see there reviews of images.

Finally, we need to label our data so that this data can be used in medical industry to develop new things.

4.2. Detailed Overview:

4.2.1. Download images from Twitter:

First we follow pathologist on twitter and then start downloading all there tweets using twitter API and extract images and its corresponding tweets. Before using Twitter API we need to construct a Twitter App. For downloading the image we use a python script, we also store the User id whose images we successfully downloaded, so that we don't again start downloading his/her images.

4.2.2. Create dataset of Pathological Images:

After downloading images from twitter we use deep learning model to filter pathological images. And then we stored filtered images in firebase storage.

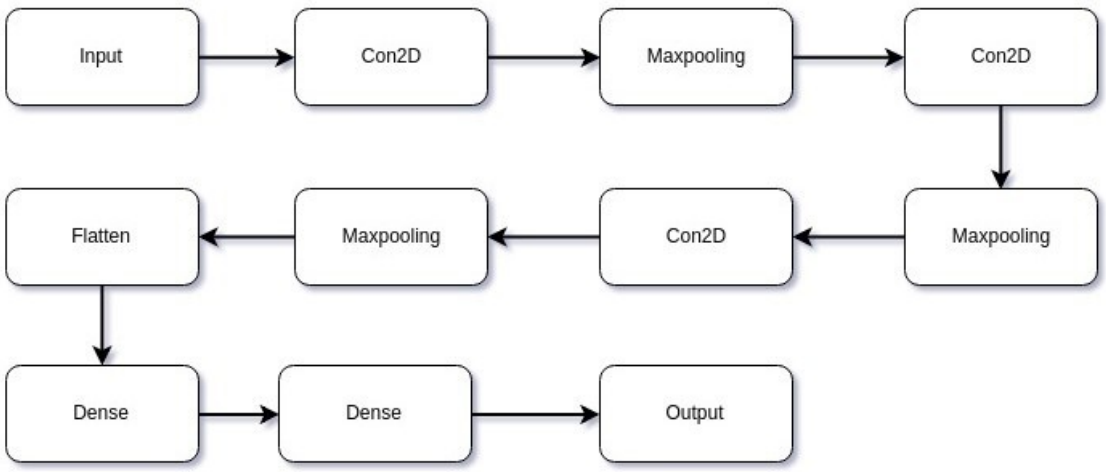

Illustration 1: Binary Classification

We use a sequential keras model to make a CNN network with three Convolution and Max Pooling layers. We then predict the output with two Fully connected layers to classify the images to be in one of the classes of pathological image or not pathological image. This is a binary classification model trained using rmsprop. We use binary crossentropy loss function to train the model.

4.2.3. Create a website:

After collecting sufficient number of images we start working on website. For website we use Vuejs platform along with Vuetify, Vuex, etc. And for database we uses Firebase database and we include various features in our website like:

Auto Login

Firebase already store token value in cache so we just use that value to again re-sign-in.

Authentication

For authentication we just read user input and then pass it to firebase inbuilt function to Signin/Signup and Logout.

Can modify self review

Only authenticated user are allow to modify self review. i.e. this component is only visible if user is authenticated. After the review is modified we again sync local database

Can change self profile

Profile page is only access by authenticated users, non-authenticated users neither can see the details nor they can modified data. And authenticated users can update only his/her own data.

Have developer's details

We created a page named 'Contact' where developer's details are given like mobile number, email, etc. This data is not locally stored but fetched from firebase database and a message button is also provided where a user can submit his/her own doubt through mail(for unauthenticated usres) and from firebase database(for authenticated users).

Can see other User's review

We provided this functionality because pathologist can be sure whether his details are correct or incorrect by comparing his details with other, and can also participate in discussion for any conflict.

About page is changed remotely

We built this page such that we can easily change its content without changing our actual file, i.e. we can set this page just using firebase database and provided this functionality so that we can run time change this page.

Can see any pathological image

Whether a user is Signin or not he/she can see our website even can see all images and all those reviews given by other users we want this feature because we want to encourage people so that more and more people become user of our website and can provide us more data.

Discussion with other Users but for authenticated users only

Its obvious that unauthorized user does not have permission to access this discussion otherwise they can create problems to us. And we want this function so that user can have discussion with each other and can help each other to make more accurate labelled data.

Search image by either index, twitter id or just type 'random' for random image

We provided these three functionality because a user want to search a random image than he can use random,

If we know the twitter id than also we can easily search image using this twitter id,

Finally an images can be search by its index in the database.

If have any problem while using website user can directly contacts developer

Either he/she can mail us or can use the 'message' button provided on the 'Contact' page of the website but be careful while clicking the button you should be login else it will redirect you to a page through which you can me us but its a tough thing for you, so it is better to prefer former one.

For every images, a hint is provided by developer itself but it may not be correct

Its just a tweets corresponding to the searched image and it may be possible that someone uploaded valid photo but use some random description for tweets then in this case provided details might be incorrect.

4.3. UML Diagram:

4.3.1. Flow Diagram

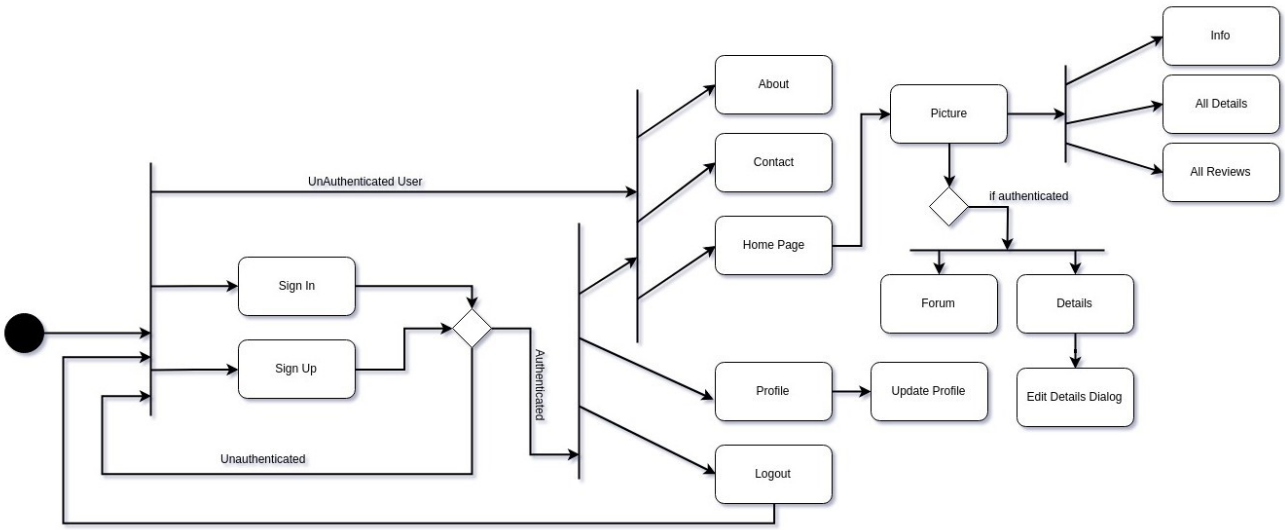

Illustration 2: Flow chart for website

4.3.2. Usecase Diagram

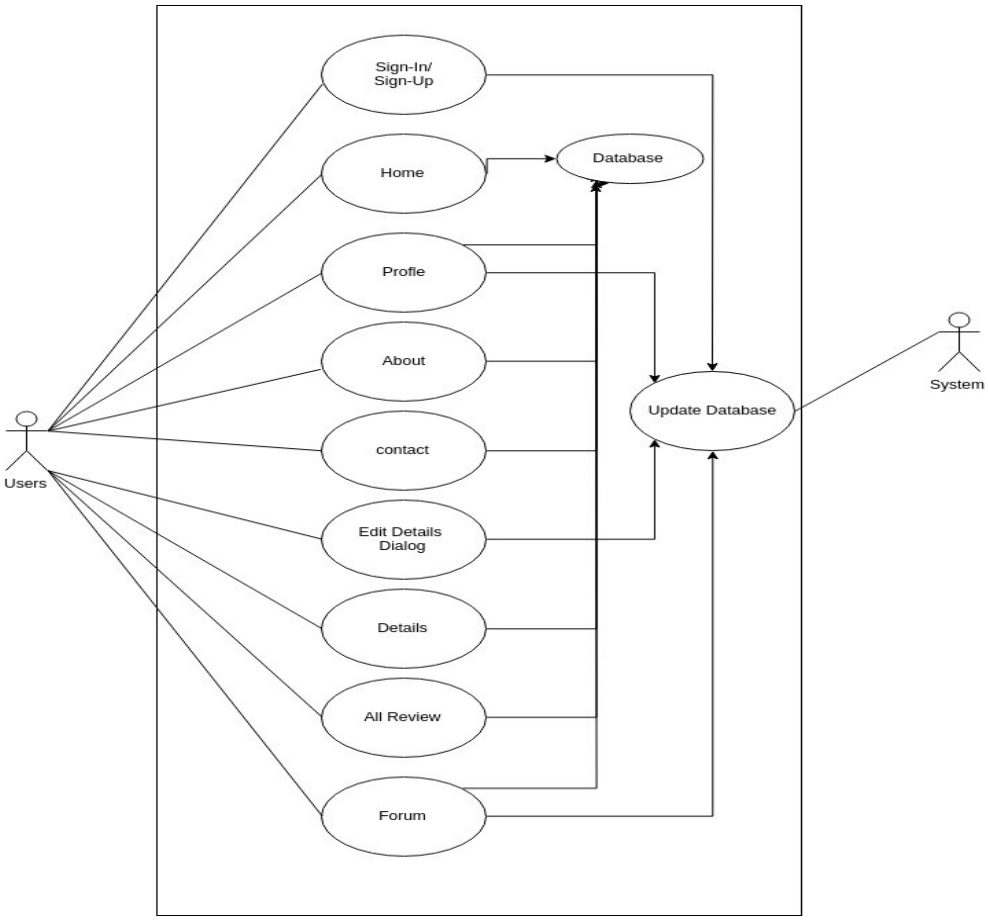

Illustration 3: Usecase Diagram

5. Future Plans

- Increase user security
- Applying ML/DL on labelled dataset for predicting the diseases, tissues, etc using images

6. Result and Conclusion

Website was made with the features like auto login, user friendly enviroment, highly secure, profile updatation, discussion form, update reviews, some backend functions of firebase, error display and and loading for removing delayed time latency.

There are some error in website like Forum message are not uploaded as expected which can be fixed in coming days.

7. References

1. https://stackoverflow.com/
2. https://vuejs.org/v2/guide/
3. https://firebase.google.com/docs/
4. https://vuetifyjs.com/en/getting-started/quick-start/
5. https://github.com/chat21/chat21-cloud-functions/tree/master/functions