caMicroscope

Proposal - Google Summer of Code 2020

Project Proposal

CONTACT INFO

- **Name:** Eman Lotfy Elrefai
- **Slack Username:** Eman Elrefai
- **Email Address:** eman.lotfy.elrefai@gmail.com
- **Telephone:** +201147520251
- **Country:** Egypt
- **Timezone:** EET
- **GitHub:** https://github.com/EmanElrefai

---

1. Project Title

Cancer Region of Interest Extraction and Machine Learning

2. Mentors

Insiyah Hajoori and Ryan Birmingham

3. Abstract

caMicroscope is a web-based image viewer optimized for large bio-medical image data viewing, with a strong emphasis on cancer pathology. This guide has sections for different kinds of use of the platform.One of the most interesting ideas created by caMicroscope is Cancer Region of Interest Extraction and Machine Learning. The aim is to support a wider range of models and flexible use of their outputs. The project has three stages: Automated Background Removal, Extracting regions of interest from images and Allow admin to limit access to specific files.

4. Code Challenge

The challenge is to make a project on interest extraction Using a machine learning toolkit of our choice. The project Was to create a tool which identifies objects in the image, Then returns positions in pixels corresponding to bounding Boxes of the object in the image.

_____L

OGS START

14th March 2020

Started my research on object detection models and studied Udemy - Machine Learning with Javascript Course.

15th March 2020

Created my custom data generator

- Collecting Images from Kaggle
- Labeling Images Using LabelImg
- The labels are in PascalVOC format. Each image has one **.xml** file that has its labels.

16th March 2020

Started detecting objects using basic yolo and shared it with mentors for feedback

17th March 2020

The feedback is migrating the code challenge to either tfjs or a webservice.Started to learn how to migrate.

19th March 2020

Created index.html and styles.css for promoting the user to choose or drop the file.

23th March 2020

built the pre-trained coco-ssd model

(coco-ssd) Object detection model that aims to localize
And identify objects in a single image.
This model detects objects defined in the COCO dataset,
Which is a large-scale object detection, segmentation,
And captioning dataset.

25th March 2020

Finally started detecting the object. My project is on CodeSandbox:https://codesandbox.io/s/object-detection-using-tensorflow-js-sw8fd

LOGS END

---

5. Project Schedule

- **Three stages**
    1. **Automated Background Removal**
    2. **Extracting regions of interest from images**
    3. **Allow admin to limit access to specific files**

Timeline

Task	Estimated Time needed for task	Estimated starting time -endingtime
Community Bonding		**May 4, 2020 - June 1,2020**
Getting familiar with all the relevant technologies/ concepts of caMicroscope. Discussing with the mentor all the details needed before beginning writing code. Establishing regular communication with a mentor.	10 days	28th April - 8th May
Development **Automated Background Removal**		**June 1, 2020 - July 3, 2020**

Doing some research on the best models for Semantic Segmentation like Unet and Tiramisu — very deep encoder-decoder architecture	3 days	
Building the selected model and getting the result	15 days	
For debugs and taking feedback .	2 days	
Checking if there are any bugs on it or exceptional situations.	5 days	
First Evaluations		**June 29, 2020 - July 3, 2020**
Development **Extracting regions of interest from images**		**July 3 - 27, 2020**
first step is detecting the edges of the segments	4 days	
Applying Morphological Transformation which involves using a succession of dilations and erosions on the image to remove unwanted edges and close gaps	3 days	
Extracting the contours in images	2 days	
Building the image mask to Allow us to pull out the desired Features from the original image.	7 days	
For debugs and taking feedback.	2 days	
Checking if there are any bugs on it or	5 days	

exceptional situations.		
Second Evaluations		**July 27 - 31, 2020**
Development **Allow admin to limit access to specific files**		**July 31, 2020 - August 24, 2020**
(View/Edit/Comment) to access all files.	3 days	
Providing a different level of access to different files.	3 days	
Last check for any kinds of problems that I will have while working	2 days	
Documentation, and trying to optimize my implementation if something is missing, and check if there is any corruptions	5 days	
Asking the community for feedback and making sure that everything works as expected, then moving to other tasks.	2 days	
Submission Code and Evaluations		**August 24 - 31, 2020**
Mentors Submit Final Evaluations		**August 31, 2020 - September 7, 2020**

6. Why this project?

I am so interested in working on Cancer Region of Interest Extraction and Machine Learning. I did a project with myself for Breast Cancer Classification. the repository of my project.
https://github.com/EmanElrefai/Breast-Cancer-Classification
My research with a professor in our university is on the same topic.It's my dream so I hope to join and work with great mentors this year.

7. Conflict of Interest or Commitment

I have no other commitments during the period of the GSoC internship.I will be having my university summer vacations from June 25th Onwards till september 24th 2020

From	To	Hours per Week
May 4	June 16	28 - 30
June 17	June 30	35 - 42
July 1	August 24	42 - 49
August 24	_____	21 - 28 (bugs and extra tasks)

8. Preferences

[1]https://medium.com/object-extraction-from-images-using-opencv-python/extracting-regions-of-interest-from-images-dacfd05a41ba
[2]https://towardsdatascience.com/background-removal-with-deep-learning-c4f2104b3157

[3]https://www.researchgate.net/figure/The-steps-involved-in-Extracting-GLCM-Features_fig3_301769862

[4]https://www.researchgate.net/figure/Various-steps-in-a-typical-region-of-interest-extraction-algorithm-a-The-filtered_fig2_221914486

[5]https://link.springer.com/chapter/10.1007/978-3-642-32384-3_1

[6]https://towardsdatascience.com/how-to-create-your-own-custom-object-detector-766cb11ccd1c

[7]https://towardsdatascience.com/detailed-tutorial-build-your-custom-real-time-object-detector-5ade1017fd2d

9. Relevant Background Experience

- Project: Breast Cancer Classification
- Project:Object Detection using basic-yolo
- Project: Object Detection using ImageAI
- Project: Object Detection with Amazon Sagemaker
- Finished Deep Learning Specialization.
- Finished Tensorflow Specialization.
- Finished Udemy- Machine Learning with Javascript Course

10. Personal

My name is Eman Elrefai. I am a third-year communication and electronics student at faculty of engineering, Alexandria University, Egypt. I am DSC Lead at Google Developer. I am a software developer and technical member at IEEE Alex sb. Successfully I completed a research internship last summer. I worked with the Natural Language Processing team and created Automatic Collecting Data and used many NLP models like TF_IDF, FastText, BERT, and BIDAF..

11. Extracurricular Studies

- Google Scholarship| Front End Track .

- information Technology institute iTi | Hello World Program : in Biblioteca Alex (Java , Data Base , Data Structure , HTML, and CSS) with final project SHAMA3A , Android app for online Shopping .
- Udacity | Developing Android Apps with Kotlin Course.
- One Milion Arab Coders Program | Android Developer Track.

12. Awards

- First place on Qodwa.Tech AI Hackathon (Reviving Culture & Heritage) powered by Microsoft , Ministry of Communications and Information Technology (MCIT), Egypt and CARE Egypt.

- First place in Alexandria Ideation Marathon Competition.
- A plus | PHD competition: Made a survey paper about Conventional neural networks .