

Proyecto procesamiento de imágenes para extracción de texto

Fecha: Miércoles 24 de enero

Hoy tuvimos la primera reunión con Paula y nos dio una introducción general a los tres proyectos.

Cosas por hacer:

- Instalar, probar y familiarizarse con el keylogger
- Leer el trabajo de Roxana (funcionamiento de keylogger) capítulo 4.
- Primer aporte generar una instalación más transparente.

Que voy hacer hoy:

- Configurar mi nueva computadora.
- Descargar pycharm.
- Clonar repositorio de Roxana.
- Tratar de instalar el keylogger.
- Familiarizarme con el keylogger.

Informe de Trabajo

- Se instalo ubuntu en la máquina que me presto Paula para desarrollar el proyecto, además configure las llaves ssh para posteriormente trabajar con git.
- Se descargo y se instalo el IDE PYCHARM ya que al ser un proyecto a desarrollar en python, esta herramienta me va ayudar a desarrollar más rápido mi código.
 - Lo descargue del siguiente link: <https://www.jetbrains.com/pycharm/>
- Clone el proyecto de Roxana del link: <https://github.com/roxana-lafuente/ResearchLogger.git>
- Crear el ambiente virtual de trabajo con los siguientes pasos:

El primer paso es definir un ambiente de trabajo basado en virtualenv:

```
sudo -H pip install virtualenv
sudo -H pip install virtualenvwrapper
```

Agregar las siguientes líneas al final del archivo ~/.bashrc:

```
export VIRTUALENVWRAPPER_PYTHON=/usr/bin/python
export VIRTUALENVWRAPPER_VIRTUALENV=/usr/local/bin/virtualenv
export WORKON_HOME=$HOME/.virtualenvs
export PROJECT_HOME=$HOME/dev
source /usr/local/bin/virtualenvwrapper.sh
Y refrescar la sesión del shell.
source ~/.bashrc
```

Para crear el ambiente virtual:

```
cd Documents/ImageKeylogger
mkvirtualenv venvImageKeylogger
```

(Para salir del ambiente utilizar deactivate #nombreEnv
y para volver a entrar workon #nombreEnv)

Generar archivo de requirements.txt:

```
pip freeze > requirements.txt
```

Fecha: Jueves 25 de enero

- Instalación de Keylogger
 - Validate [sudo pip install validate]
 - Xlib [sudo apt-get install python-xlib]
 - Python Image Library (PIL) [sudo apt-get install python-pil]
 - gtk [sudo apt-get install python-gtk2]
 - Tkinter [sudo apt-get install python-tk]
- Instalación más fácil para Linux, cree un script para que la instalación en linux sea mucho más fácil y sencilla para los usuarios.

```
#!/bin/bash
```

```
#start
```

```
echo Starting the installation of Keylogger
```

```
#update package  
sudo apt-get update
```

```
#install necessary software  
sudo apt-get -y install python-pip python-xlib python-pil python-gtk2 pytho  
sudo pip install validate
```

```
#finish  
echo Installation completed
```

- Probar el script. Para probar el script voy a hacer una máquina virtual con ubuntu 14.04 para verificar que el script sirva desde esa versión hasta la más reciente.
 - Descargar virtual box: wget http://download.virtualbox.org/virtualbox/5.1.16/virtualbox-5.1_5.1.16-113841UbuntuYakkety_amd64.deb
 - Luego instalar el paquete: sudo dpkg -i virtualbox-5.1_5.1.16-113841UbuntuYakkety_amd64.deb
 - Descargar ubuntu 16.04: <https://www.ubuntu.com/download/desktop>

Fecha: Viernes 26 de enero

- Se asignó la lectura “Keylogger para el estudio de los procesos cognitivos del traductor”
 - Leí del libro los primeros dos capítulos y los anexos técnicos.

Pruebas

- Empezar hacer pruebas de correr el programa [ResearchLoggerInterface.py](#) y [ResearchLogger.py](#)
 - Tengo problemas a la hora de terminar el programa con Ctrl izq + Ctrl der + F12
 - Estoy analizando el archivo [onclickimagecapture.py](#) ya que es acá donde se toman las capturas de pantalla que más adelante voy a necesitar.
- Intenté hacer una máquina de virtualbox en mi computadora pero no me sirvió debido a que no permite la virtualización.
- Hice una máquina para probar que el instalador de Linux sirve perfectamente en versiones mayores a 14.04.

- Baje el instalador el la máquina de test.
- Clone el proyecto en la máquina test.
- El instalador funciona correctamente.
- Actualice la bitácora.

Fecha: Domingo 28 de enero

Nuevas sugerencias

- Arreglar script para que valide si fue instalado o no.
- Arreglar dependencias
- Que descargue el repositorio de ResearchLogger
- Investigar como se puede a través de bash poder validar que un paquete este o no instalado en la computadora.
 - Encontré que con la funcionalidad de “dpkg” se podría realizar una consulta al sistema y saber si este estaba o no instalado.
 - Para solucionar el problema de dependencias lo que hago es tratar de arreglar el sistema en caso de que ya tuviera errores con las dependencias y luego aplicó una actualización de paquetes para que no ocurra ningún error a la hora de instalar.
- Investigar si se podría validar un paquete instalado por pip
 - No encontré la forma intente con \$(pip freeze | grep validate) pero en mi computadora aparece como si no estuviera dentro de los paquetes instalados pero si lo tiene.
 - Para minimizar el problema de arriba lo que hice fue ya que como utilizo “pip” para instalar el paquete “validate” lo que hago es actualizar pip para que no haya ningún error cuando instale el nuevo paquete.
- Mostar una salida donde quede más claro que fue lo que ocurrió.
 - Necesitaba colocar salidas de warning o de su correcto funcionamiento por lo que la salida se imprime de color rojo cuando es un warning y verde cuando su funcionamiento es correcto.

Fecha: Lunes 29 de enero

- Tarea de iniciar con la herramienta Tesseract utilice los siguientes links para iniciar
 - <https://www.pyimagesearch.com/2017/07/10/using-tesseract-ocr-python/>

- <https://www.pyimagesearch.com/2017/07/03/installing-tesseract-for-ocr/>
- <https://www.pyimagesearch.com/wp-content/uploads/2014/02/ImageSearchEngineResourceGuide.pdf>

- Instalar tesseract

Primero instale los paquetes

```
pip install pillow
pip install pytesseract
pip install cv2
pip install opencv-python
```

Tuve problemas lo solucione con crear un ambiente virtual, lo que pasaba era que pillow tenia una doble referencia por lo que se necesita hacer el ambiente virtual y que lo único que este dentro de este ambiente sea este paquete.

```
pip install pillow
pip install pytesseract
pip install cv2
pip install opencv-python
```

- Pruebas con la herramienta tesseract-ocr

- Primero inicie con la imagen de prueba con letras grandes y la herramienta trabaja muy bien, el resultado fue excelente.
- Luego utilice una imagen "salt and pepper" osea que tenga ruido en el fondo de la imagen, esto con el fin de ver si se tiene el mismo resultado.
- La siguiente imagen era un código el cual tenia letras de colores y a la hora de aplicar la herramienta esta borra las palabras de color.
- Hice pruebas con las imágenes de los LOGS del ResearchLogger al pasarlo por la herramienta el resultado no es para nada bueno, fue donde note que existe una necesidad de una buena resolución.

- Mejorar Resolución

- Para observar la resolución de la imágenes en el servidor utilice:

```
identify -ping image.png
```

- Resolución 150*150, este tipo de imagen carece de información en la resolución, al probar este tipo de imagen con la herramienta se pierde

información al extraer el texto dando como resultado erróneo.

- Resolución 350*350
- Resolución 500*250,

Nota:

Pienso que no del todo es responsabilidad de la resolución también la herramienta tiene dificultades al ser imágenes con colores. Para que el resultado sea bueno debería de tener fondo blanco y letras negras.

Fecha: Martes 30 de enero

1. hice el archivo [ocr.py](#).

- Este archivo lo que hace es procesar un archivo con la herramienta tesseract
- Tiene tres formas de correr:
 - python [ocr.py](#) --image example_01.png
 - python [ocr.py](#) --image example_01.png --preprocess thresh
 - python [ocr.py](#) --image example_01.png --preprocess blur

2. Trate de correr el ejemplo de pochetti para tratar de encontrar respuestas.

- Puedo correr solo la mitad del proyecto de pochetti pero me da la parte que necesitaba, la parte del reconocimiento de objetos dentro de la imagen. Con esto puedo darme una idea de como es que se realiza este proceso.

3. Leí capítulo 3 y 4 keylogger para los procesos del traductor.

Fecha: Miércoles 31 de enero

1. Reunión con Mariona de Lérida para la instalación de ResearchLogger

2. Necesito parsear el Log de los clicks.

- Lo que hice fue hacerlo con bash.

```
cat /home/tito/Documents/ResearchLogger/logs-tito-1517412618/clic
```

- Estoy sacando:

Xdown,Ydown, Nombre de la imagen, Xup, Yup
30,58,20180131_123027_241486_gnome-terminal-server_tito.png,30,58

3. Al parsear el archivo pretendo hacer:

- Sacar el texto de las imágenes.
- Verificar donde dio click dentro de la imagen.
- Verificar si arrastro un click (con esto pretendo saber si subrayo algo o no)

4. Solo hice lo de Verificar donde dio click dentro de la imagen.

Fecha: Jueves 1 de febrero

1. Estoy pensando si es necesario o no sacar también a la hora de parsear:

```
Xdown,Ydown, evento, Nombre de la imagen, Xup, Yup, evento  
30,58,left_down,20180131_123027_241486_gnome-terminal-server_tito.png,  
  
cat /home/tito/Documents/ResearchLogger/logs-tito-1517412618/click_ima
```

2. Voy a verificar si arrastro un click

- La idea es correr el ResearchLogger y hacer algunas pruebas subrayando algún texto. Teniendo esas pruebas se le corre el software [coordenadas.py](#) que lo que hace es leer el Log de los clicks y saca la información de las coordenadas para poder identificarla en las imágenes. Con esto vamos a saber si las coordenadas son distintas de los dos diferentes eventos(presionar click, soltar click).

3. Paula me paso para que leyera este link:

<https://www.pyimagesearch.com/2017/06/19/image-difference-with-opencv-and-python/>

Fecha: Viernes 2 de febrero

1. Reunión con todos.

Fecha: Lunes 5 de febrero

1. Estoy haciendo el fork del proyecto y ordenando las cosas del repositorio anterior(imageKeylogger)
2. Haciendo pruebas con la biblioteca de pynput para tratar de loguear todos los movimientos posibles y clicks del mouse.
Esta herramienta utiliza las funciones:
on_click: esta función lo que hace es obtener todos los eventos relacionados con la captura de los botones del mouse, se puede saber cuando presiona una tecla y suelta una tecla. tiene las 3 teclas(click izquierdo,derecho y centro)

on_move: esta función logra captar todos los movimientos del mouse a lo largo de la pantalla, por ejemplo si mueve el mouse sobre una palabra o si no lo esta moviendo.

on_scroll: es la función que obtiene el scroll hacia arriba y scroll hacia abajo con esto podemos saber si se encuentra en un documento analizando ciertos detalles o navegando por la web.
3. Pruebas para definir el cropbox. el tamaño de la ventana.
Se necesita calcular un tamaño adecuado o un valor por default que con base en algo podamos determinar cual es el tamaño ideal que debería de tener la imágenes que se toman para que el post análisis sea más fácil y el resultado sea lo más exacto posible.
4. pyscreenshot es una biblioteca que toma tanto captura completa, pero también cuenta con la funcionalidad de poder tomar solo un pedacito(cropbox).

Fecha: Martes 6 de febrero

1. Hacer un script que capture cualquier movimiento con el mouse.
Para realizar este script utilice la documentación que tiene la biblioteca <https://pypi.python.org/pypi/pynput> para familiarizarme con la herramienta utilizó los ejemplos por defecto de la herramienta.
2. Voy a hacer el mismo script de prueba para windows(con el fin de ver si al cambiar el código tenemos la compatibilidad)
 - Problemas con el click derecho, cada vez que se utiliza ese click se detiene y ocurre un error

- El error se solucionó determinando que era solo si el click derecho se realiza sobre la terminal de windows donde se ejecuta, si es en cualquier otro lugar funciona bien.

3. Realizar cambios al anteproyecto para que quedé listo para mañana.

Fecha: Miércoles 7 de febrero

1. Empece a realizar cambios en el proyecto de ReseachLogger.

Necesito arreglar necesariamente el como se toman las imágenes de la herramienta para que me sea más fácil la captura de elementos dentro de las imágenes por lo que voy a hacer es ver el código por encima para ver que funcionalidades son necesarias y cuales no.

- Empece con las constantes
- La funcionalidad de agregar scroll podría ser interesante.
- Existe posibilidad de agregar ruta donde el usuario este trabajando.
- Existe posibilidad de agregar status del proceso, ej: running

Empece a ver todo el código y me esta costando entender algunas cosas. tiene muchas clases dentro de los archivos y hay muchas funciones repetidas dentro del código, me estoy frustrando un poco, estoy viendo otras posibilidades. Estoy llegando a pensar que tal vez es mejor empezar hacerla desde cero.

2. Reunion de seguimiento con Paula.

Fecha: Jueves 8 de febrero

1. Pienso en hacer otro ResearchLogger con solo la biblioteca pynput.

2. Realice un mapeo de todos los archivos que debo tocar para realizar los cambios.

- [constantes.py](#)
- [detailLogWritter.py](#)
- [onClickImageCapture.py](#)
- [pyxhook.py](#)
- [ResearchLogger.py](#)
- [timedScreenshot.py](#)

3. No se que hacer pero voy a seguir.

Fecha: Viernes 9 de febrero

1. Reunión de seguimiento con Aurelio
2. Reunión con Paula.

Lo que voy a empezar hacer es hacer pequeños cambios de algunas funcionalidades y ir testando si todavía a pesar de los cambios la herramienta Reseach logger sigue funcionando con normalidad.

3. Empece realizando cambios para obtener el process_name y verifique los logs para ver si todo sigue bien.
4. Lo del process_name no anda bien. Lo estoy probando dentro de la herramienta pero a la hora de verificar en el log no se esta mostrando nada.

Fecha: Lunes 12 y martes 13 de febrero

FERIADO

Fecha: Miércoles 14 de febrero

1. Estoy realizando un documento que resume lo que he hecho por semana en las semanas anteriores.
2. Reunion con Aurelio.
3. La reunión me aclaro mucho las ideas ya que me volvió a poner en claro que es lo que debo hacer y cual es mi objetivo en este proyecto.

Fecha: Jueves 15 de febrero

Hoy voy a testear la toma de screenshots en la Herramienta

Cambie la biblioteca de tomar los screenshot por pyautogui uso lo mismo que la biblioteca anterior pero esta tiene una mayor compatibilidad con la biblioteca cv2 por lo que me va ayudar más cuando tenga que hacer procesamiento de imágenes.

Fecha: Viernes 16 de febrero

1. Estoy tratando de asegurar de que todas las imágenes que se tomen, su respectivo x,y quede dentro del cropbox. El asegurarlo se me esta haciendo difícil ya que depende del lugar donde se tome queda fuera de la pantalla. Quiero no cambiar

todo el código que ya se tiene para esta funcionalidad trató de reutilizarlo.
Voy a tratar de buscar una biblioteca que lo haga.

2. Cree el archivo [saveimage.py](#)

La idea detrás de este script es poder hacer un caso de prueba donde yo pueda tener a partir de un x,y que yo establezca tomar el respectivo screenshot y un cropbox del mismo tratando de simular lo que ocurre en la herramienta pero sin tener que correr las otras funcionalidades de la herramienta y otros errores que no nos interesan.

Fecha: Lunes 19 de febrero

1. Hoy voy a cambiar para que las capturas completas de pantalla se hagan con la nueva biblioteca.
2. Las capturas de pantalla(pantalla total) están listas.
3. Los procesos no se van a poder cambiar a este formato, El problema es que el evento no trae el id o el proceso actual por lo que no encuentro una manera de hacer este proceso genérico para que funcione multiplataforma.

```
p = psutil.Process(os.getpid())  
p.name()
```

Fecha: Martes 20 de febrero

1. Hoy termine de modificar todas las funciones de las imágenes, las imágenes ya están perfectas para empezar analizarlas.
2. Realice modificaciones al [saveimage.py](#)

Voy a dejar hasta acá los cambios del research logger y voy a empezar con el análisis de las imágenes porque siento que me estoy atrasando un poco. si queda tiempo realizo los respectivos cambios.

Fecha: Miércoles 21 de febrero

1. Reunión con Paula y con los demás compañeros analizamos en conjunto los problemas que teníamos cada uno y nos hacíamos sugerencias.

2. Tuve una conversación con Paula y le dije sobre que me iba a enfocar en lo de las imágenes, que iba a dejar un poco de lado lo de agregar nuevas funcionalidades y fijamos un nuevo norte.

Hablamos sobre los paper y dijimos lo que pensábamos hacer en las próximas semanas ya que Paula no va a estar en el país.

NECESITO SABER EN QUE ÁMBITO PUEDE SER BUENO ENFOCAR MI INVESTIGACIÓN

* Buscar plagio

3. Le dije a paula que necesitaba ir conociendo las actividades que realizaba Laura para sacar información observando las imágenes.
 - Me dio un borrador de la tesis que esta realizando Laura para ser master en traductología

Fecha: Jueves 22 de febrero

1. Hoy me voy a dedicar a leer la tesis de Laura
2. Cosas importantes de la tesis
 - Es muy importante el uso de la herramienta OmegaT
 - cuenta con 10 pasos que todo traductor necesita hacer dentro de esta herramienta se podría utilizar el análisis de imágenes para reconocer estos pasos.
 - Reconocer cuando trabaja en LibreOffice writer
 - Reconocimiento de archivos (puede ser también la extensión ej: .png, .jpg, etc.)
 - Es importante los 18 screenshots de análisis
 - Lograr saber cuando un usuario se da cuenta de un error y lo corrige
 - Encontrar segmentos sin traducir dentro de la herramienta (figura 31 muestra 2)
 - Verificación de etiquetas (vendría siendo parte de los pasos)
 - Crear documentos finales
 - Activación del glosario (screenshot)
 - Activar coincidencias parciales (screenshot)
 - Activar Diccionario (screenshot)
 - Subrayado de texto
3. Le ayude con algunas funcionalidades a Walter

- Verificar que las flechas funcionen correctamente en Linux
- Que el “tab” se captado por la herramienta

Fecha: viernes 23 de febrero

Detección de objetos

Existen tres procedimientos principales en el análisis de BLOB:

1. Extracción: En esta parte del procesamiento se encuentran los pixeles conectados entre si o sea pixeles vecinos que coinciden con las especificaciones, para lograrlo usualmente se utilizan algoritmos como el algoritmo Grass-Fire Recursivo o el algoritmo Grass-Fire Secuencial
2. Representación: Es la de representar las características de los pixeles que hemos extraído anteriormente, esto se hace por medio de números que indica la cantidad de características y luego puede ser comparado por medio de un método para asegurarse de que cumple con las características propuestas, por ejemplo si es un rectángulo, los pixeles deben representar cuatro lineas interconectadas, dos paralelas entre si horizontalmente y dos paralelas entre si verticalmente.
3. Clasificación: Se compara directamente con un prototipo para verificar que realmente concuerda con las características que queremos filtrar, por ejemplo en el caso anterior compararíamos los pixeles con un conjunto de pixeles que representan un rectángulo.

<https://www.pyimagesearch.com/2015/09/07/blur-detection-with-opencv/>

<https://planetachatbot.com/introducción-a-la-detección-de-objetos-con-python-y-simplecv-ccbbe33a5bc4>

Estoy realizando pruebas para entender como puedo solucionar el problema de seleccionar el centro como región de interés y si a la hora de analizar toda la imagen se puede seleccionar que analice solo una parte de la imagen.

Fecha: Lunes 26 de febrero

Hoy estoy tratando de resolver como identificar que dio click sobre un menú y sobre que parte del menú hizo click.

1. Estoy realizando el script textDetect.

- Quiero intentar detectar solo los objetos del centro de la imagen
- Estoy realizando diferentes funciones con diferentes contornos y kernels para poder observar cual me sirve mejor.

2. Puedo tener la región de interés con:

`cv2.selectROI(im) # esta debe ser con un humano`

3. ya puedo tener todos los objetos que hay en una imagen

Fecha: Martes 27 de febrero

Desde el día de ayer estuve viendo como enfocarme en la parte central de la imagen, ya que como es acá donde esta el click es la zona donde voy analizar para identificar que voy hacer.

* Estaba realizando una extracción de la parte central(forma de rectángulo) pero esta no me servia me parecía extraño que me estuviera ocurriendo eso.

El problema era el siguiente:

En la función yo me encontraba realizando el orden de parametros de esta forma

`(x1,y1),(x2,y2)`

pero realmente era

`(y1:y2),(x1:x2)`

El orden cambiaba... creo que estaba demasiado estresado... :(

Mi otra teoría de hoy es comprobar si se puede sacar de una imagen por algún color en específico.

* Observe que esta teoria no me funcionaba si lo hacia con el color que yo necesitaba. Creo que solo sirve con el color rojo,azul,amarillo y gris.

Fecha: Miércoles 28 de febrero

1. Posibles Soluciones:

- Sikuli <http://www.sikuli.org/>
- Scikit-image http://scikit-image.org/docs/dev/auto_examples/features_detection/plot_blob.html
- Blob análisis:
 - Bounding box of a blob is the minimum

- object detection using blob tracing

2. Reunión con Aurelio.

Fecha: Jueves 1 de marzo

1. En la reunión de ayer con Aurelio me dijo que era muy necesario realizar mis informes de avances por lo que hoy decidí empezar hacer todos los informes desde que inicie mi proyecto.
2. Realizó algunos cambios en mi bitácora.
3. Voy hacer una carpeta en google drive donde voy a ir haciendo cada informe de avance.