**How to install LabelImg in Windows with Anaconda?**

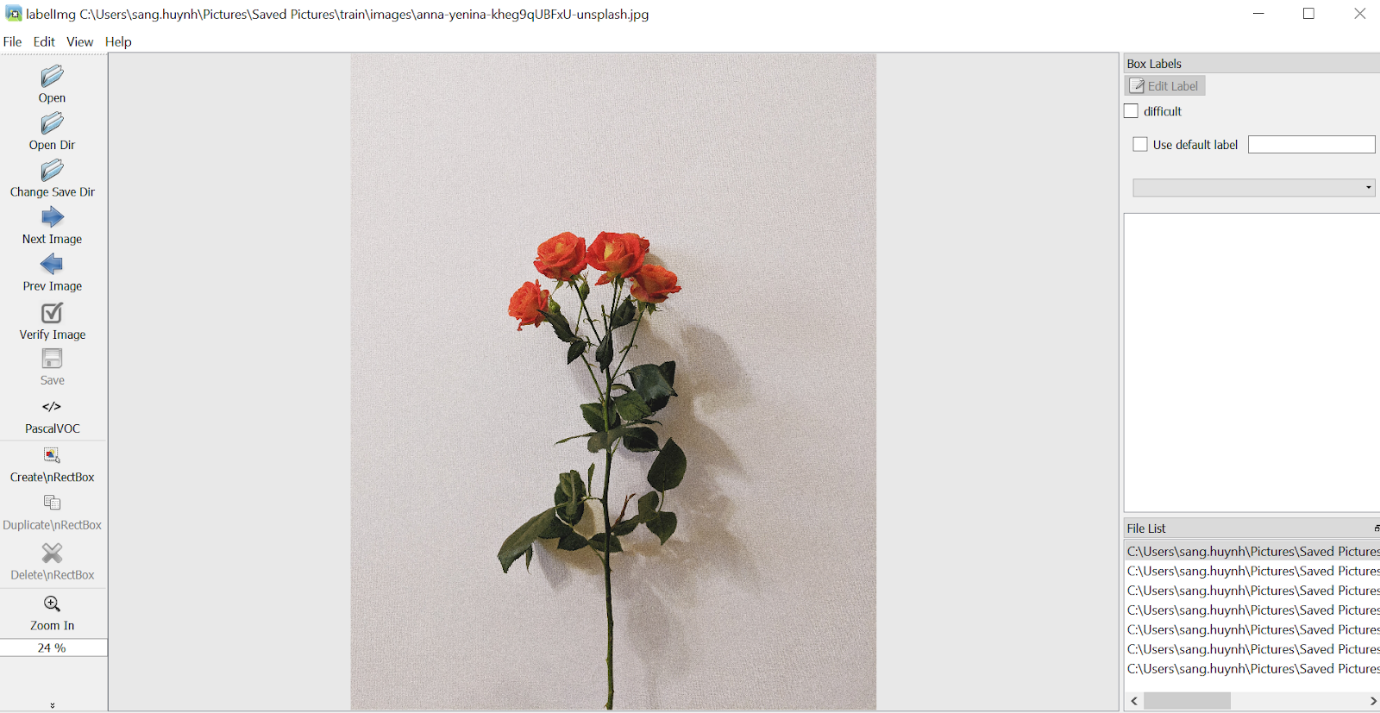
**1. What is Data labeling?**

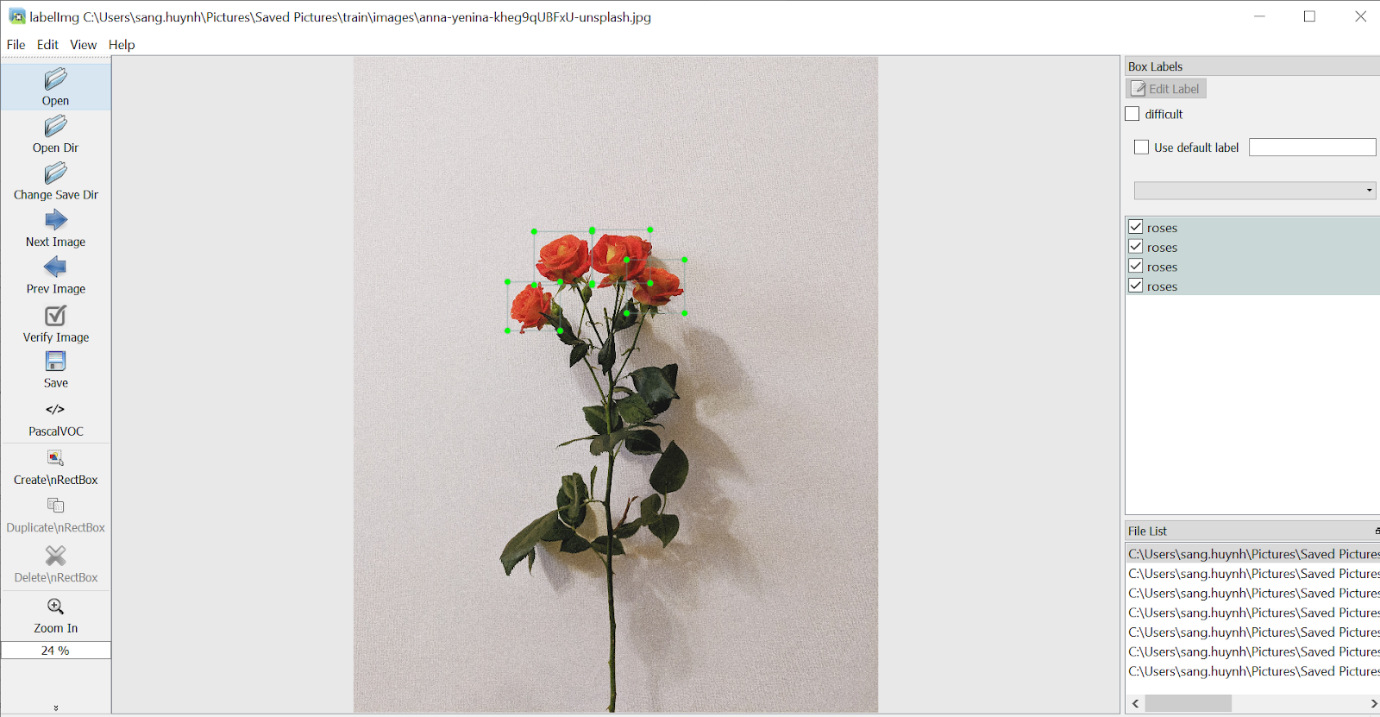
One of the most important part in the data science process is data pre-processing which makes data scientists put lots of time and effort into. Particularly, in the context of machine learning, data pre-processing requires a step of labeling. This is the step to detect and label data sample into multi-classifications so that the labeled data can be used in further machine learning process. Two typical labeling data type in ML projects are text and image annotation. The former is the crucial step in NLP projects such as sentiment analysis and the latter is the required part in computer vision projects. Labeling process is normally handled by manual, but it can be assisted by some applications to reduce the working-time. In this article, I will introduce **LabelImg**— an image annotation tool which is quite simple and easy-to-use for the beginner or non-technical guys who are concerning about how to do computer vision projects.

**2. What is LabelImg?**

**LabelImg**is a graphical image annotation tool which provides images with bounding boxes after labeling. It is written in Python and uses Qt (one of the most common GUI for python) for its graphical interface. The output labeled data are saved as XML files in PASCAL VOC format, the format used by ImageNet. Also, YOLO format is supported in this tool.

Below are the LabelImg User Interfaces when opening:





**3. Why do we use LabelImg?**

* It is open-source then it is free to use.
* It has both online and offline versions. The offline interface is user-friendly and allows the user to make fast annotations.
* The output annotation images are saved separately as XML files along with Pascal VOC format or YOLO format. These are common formats supported by many ML packages.
* After installation, it is easy to use with simple instructions.
* It can process a large number of images when compared to other free apps such as **Cloud Vision API** (the free-offline version is limited up to only 2000 image files) or **Labelbox**(the free version only allows maximum of 5000 images to be processed).
* It can be run in multiple operating systems such as Linux, Mac or Windows.
* It is written in Python then it can interact with other python packages and is easy to be used in python applications for further development.

However, this tool still has some limitations:

* Bounding box (rectangle shape) is the only region shape to be used. Therefore, it cannot give an exactly detection for complicated objects with curved lines…
* The installation requires some tricks we need to know. Therefore, I will present how to install this tool step by step in the next session.

**3. Installation.**

**3.1. Prerequisites.**

The installation below is used for **Windows OS**with **Anaconda**. Therefore, we need to install **Anaconda** to have an **Anaconda terminal**.

Follow the link below to install **Anaconda** in Windows:

<https://docs.anaconda.com/anaconda/install/windows/>

**3.2. Installing LabelImg in Windows with Anaconda.**

* Install **Anaconda** as described in the above session.
* Create a new environment in **Anaconda**.

We should create a new environment instead of using the base environment in **Anaconda** to avoid conflict of versions while installing the packages. Because the current **LabelImg**packages requires the older version of python (< 3.8.x). Therefore, creating a new separate environment can avoid the error conflict when there are two versions of python in a same environment.

Do the following commands:

conda create -n env\_name  
conda activate env\_name

E.g:

conda create -n labelimg  
conda activate labelimg

**3.3. Install python/tensorflow.**

Please note that we only install the python version which is older than 3.8.x because the latest tensorflow version is not compatible with python version > 3.8.0.

Do the following commands:

* Install python (version < 3.8.x):

conda install python=3.7.0

* Install tensorflow:

conda install tensorflow  
# this command will install the latest version of tensorflow

**3.4. Install pyqt and lxml.**

These are prerequisite packages needed to be imported before running labelImg.

* Install pyqt 5:

conda install pyqt=5

* Install lxml:

conda install -c anaconda lxml

**3.5. Install labelImg package.**

* Download package github labelImg from <https://github.com/tzutalin/labelImg> and extract into folder.
* Change the current directory to the folder of **labelImg.**

cd **path\_to\_lablelImg\_folder**

e.g:

cd C:\Users\sang.huynh\Documents\SANG\labelImg

* Do the following command:

pyrcc5 -o resources.py resources.qrc

This command will create a **resource.py**file in folder **labelImg**. This file will be moved to the libs folder in the next step.

* Change the location of **resources.qrc** and **resource files** to **libs**folder(libs folder is in labelImg folder). Please note that this step is very important to have a successful installation of **labelImg**.
* Finally, do the below command:

python labelImg.py

Then it opens a window of **labelImg** app. In View menu, tick Auto Save Mode — this will make the output files are automatically saved in Dir Folder. Output files are saved in XML format as default.

**4. Open labelImg after the first installation.**

* Open anaconda terminal;
* Change the current directory to the**labelImg** folder (please remember to change the current directory path in anaconda to the **labelImg** folder when opening this tool):

cd **path\_to\_lablelImg\_folder**

* Do the follow command:

python labelImg.py