DDSM Utility v3.1

Anmol Sharma
Undergraduate Engineering Student
DAVIET Jalandhar

January 21, 2016

Contents

1 Introduction

This is a tutorial for the DDSM Utility to greatly simplify downloading, converting, viewing and extracting annotations from the Digital Database for Screening Mammography (DDSM) database available here

DDSM is a very famous mammogram database which researchers around the world use to test their abnormality detection, and classification algorithms so as to ensure that the results are comparable to other authors or systems. However, an open source utility or tool doesnt seem to exist to easily and efficiently download the database. The website where the database is hosted has not been updated from the last 15 (almost 16) years, and hence the deprecated and obsolete software that comes with it is essentially, unusable. None of the authors/labs/universities using the database has released open source code that can process DDSM images.

The utility is capable of automatically parsing ICS files which contains the LJPEG image sizes to allow LJPEG to LJPEG1 and further to RAW decompression, and conversion into other known formats like PNG. I decided to write this utility when I was looking for ways to download and convert DDSM dataset, only to find out that the images are in a very old LJPEG format. The only utility available on the DDSM website (written for very old SunOS 5.6) was to convert it into LJPEG1 (another old, obsolete and unusable format), and that too was highly labor intensive because it required the user to manually run the script for every image he wanted to convert. Even if the user did that, converting LJPEG1 file to PNG was a nightmare, if not impossible.

Another utility available was written by Dr. Chris Rose of University of Manchester(here). It required the user to input names of IMAGE files manually, which then his script used to download the DDSM image in PNG format. This was again, ridiculously labor intensive, and didn't make sense if the user wants to download, say, 1000 images. The source code was not supplied, so making changes to the internal working was not possible. Moreover, if the user wanted to download the annotation files along with the image, he has to repeat the steps all over again. I wasn't getting paid by the hour, so this was downright ridiculous for me.

2 Contents

openDDSMLJPEGAndConvertToPNG.m – Main MATLAB script.
ConvertDDSMImageToRaw.m – Supporting MATLAB script to open the compressed LJPEG files and then convert the files into PNG or any other user defined format. Uses the utility jpeg and ddsm2raw.exe files written by Dr. Chris.
openDDSMPngWithOverlay.m – MATLAB script to open the converted PNG files, read their corresponding OVERLAY files, read their annotations and display the boundary details by superimposing on the mammogram for simple viewing.
readBoundary.m – Supporting MATLAB script to read OVERLAY files (Original author Dr. Jayasree Chakraborty, Memorial Sloan Kettering Cancer Center, New York.)
Jpeg.exe - Windows executable file to convert LJPEG to LJEPG1 (original author Dr. Chris Rose, University of Manchester)
ddsmraw2pnm.exe - - Windows executable file to convert LJPEG1 to PNG or any other format (original author Dr. Chris Rose, University of Manchester)

3 Features

- Easily download multiple case files automatically using WinSCP (Not included)
- Normalize LJPEG images according to their scanners automatically! No manual normalization needed.
- Automatically convert LJPEG DDSM images into PNG, JPEG, TIF, GIF and other formats.
- Read boundary information of each of the converted images and display them.

4 Requirements

- A Windows based machine.
- WinSCP
- Cygwin (32 BIT VERSION ONLY)
- MATLAB 201x

Tested On:

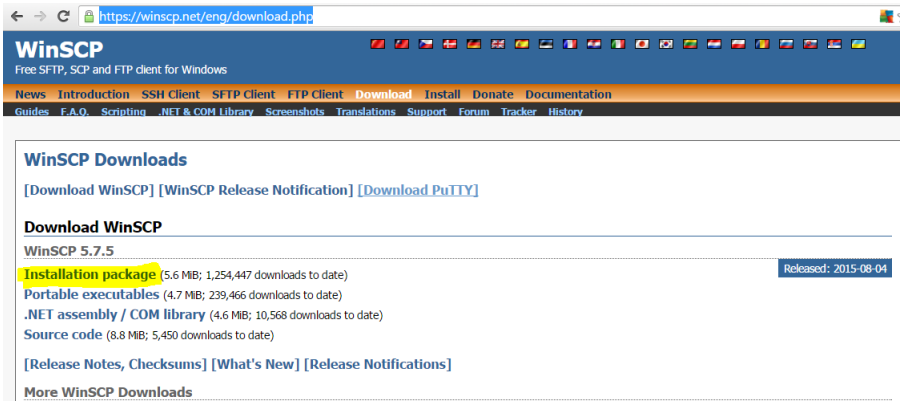
- Dell Inspiron 15 3537
- Intel i5 4200U
- 4GB RAM
- Cygwin 32 bit with ImageMagick and Ruby Interpreter.
- Windows 10 with MATLAB R2015b
- DDSM Software by Dr. Chris Rose.

5 Steps

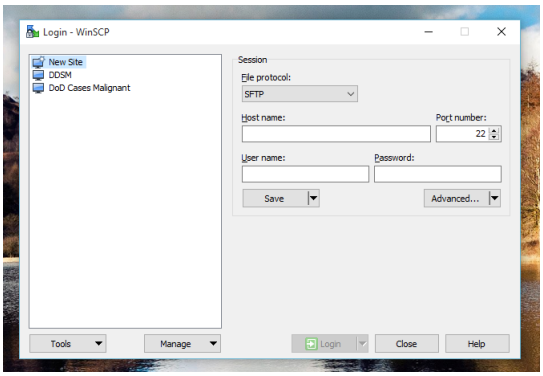
Follow the instructions and click on yellow highlighted regions in the figures. I suggest that you very carefully go through the steps, and not miss anything. Please also make sure that you perform each step in order.

- Download the required DDSM cases using the FTP link from USF.

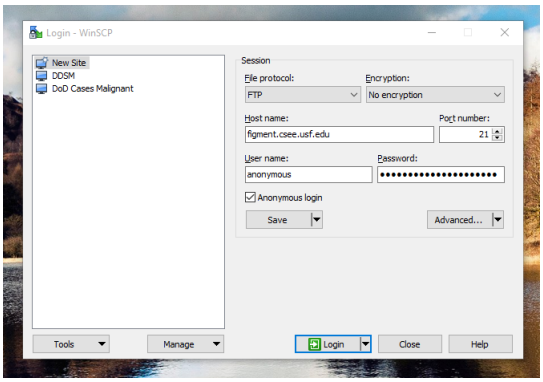
- Download WinSCP (here).
- Download the latest version as shown in figure.



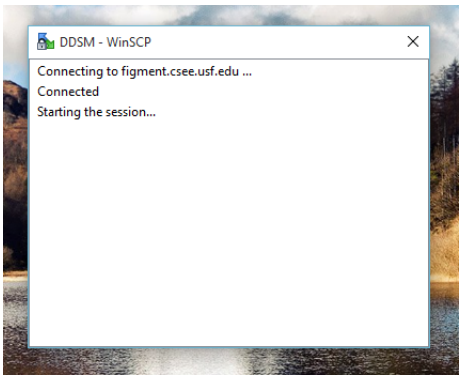
- Install the software, and open it.



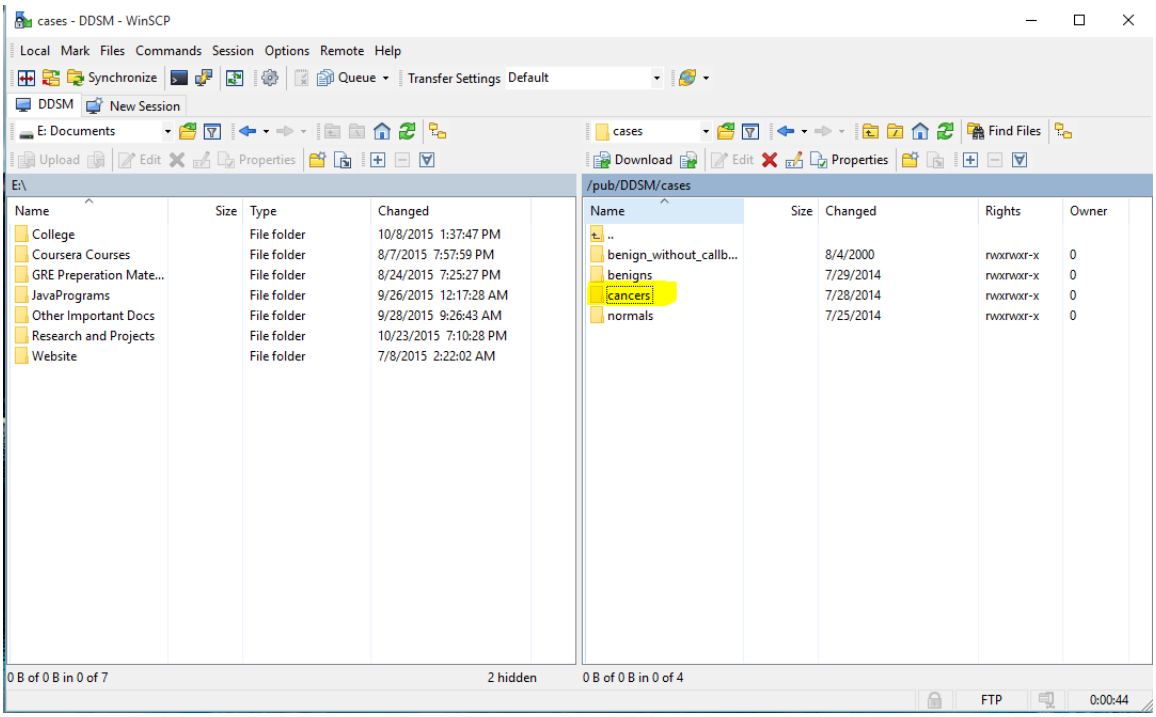
- WinSCP will ask you for the website and login details for the FTP server. Enter the credentials as shown below.

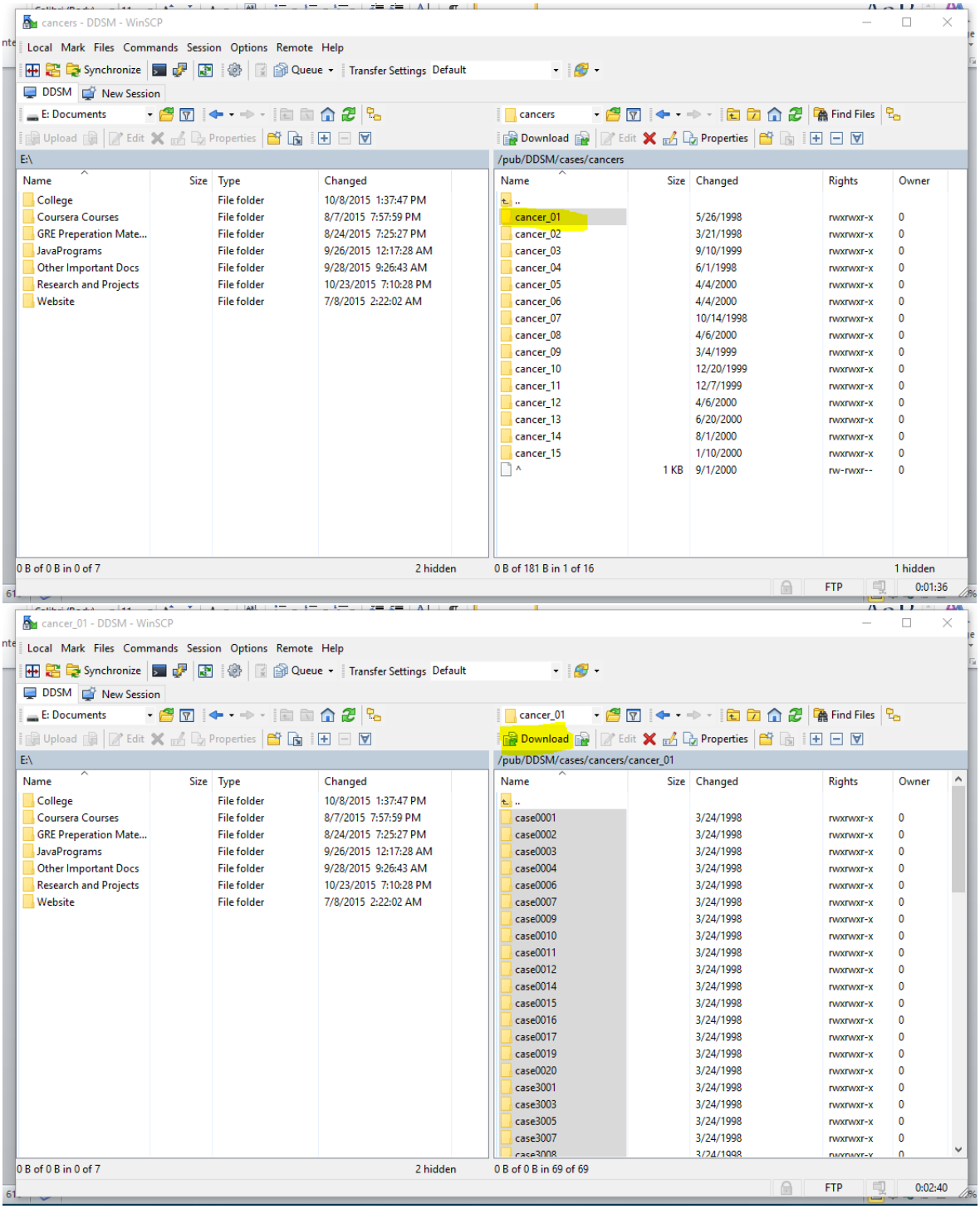


- WinSCP will connect to the server.



- Once connected, you will get the following screen. On the right, you have the FTP server parent directory, and on the left you have your own local computers directory. Now you just need to select the files (cases) you want to download from USFs FTP server by selecting them from the right side, and then click on download button on the right to start the download to the left pane location.

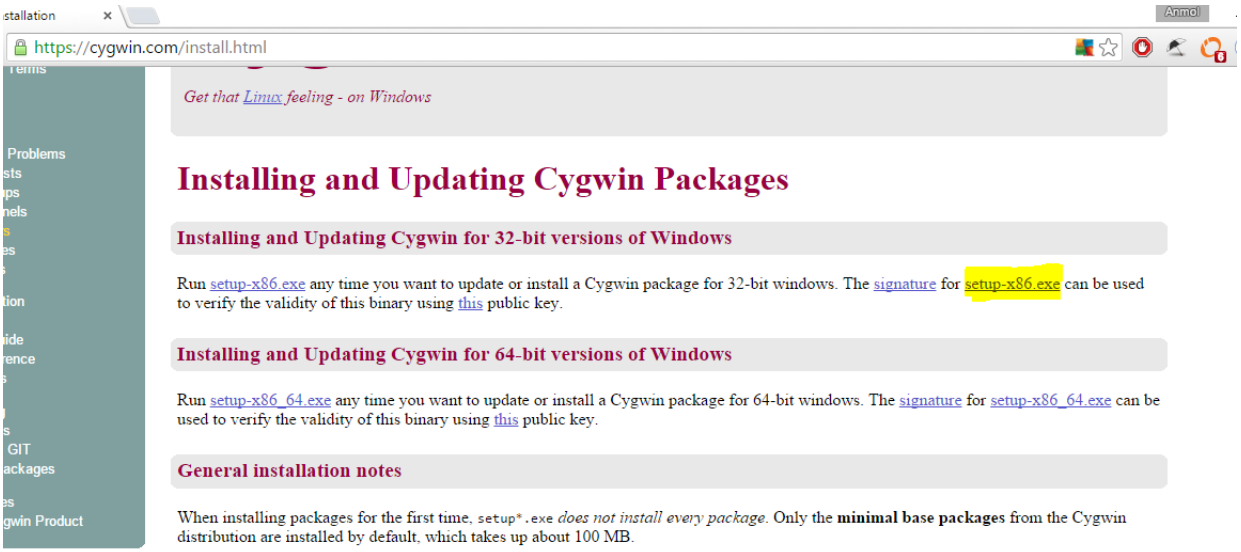


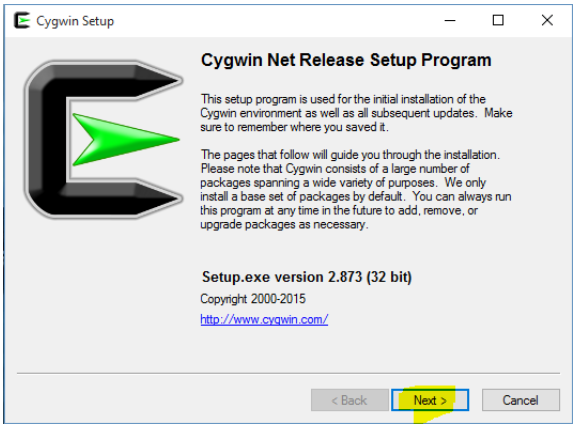


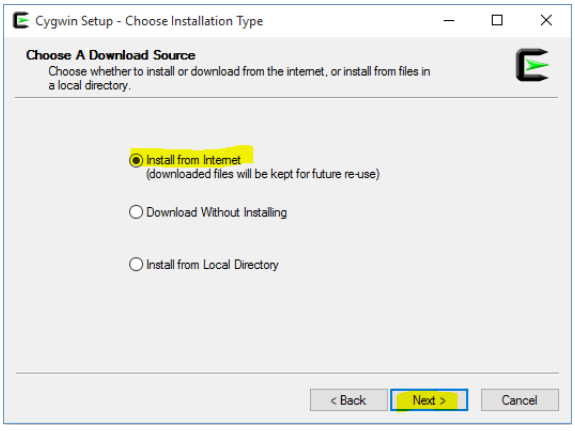
- I pressed CTRL+A to select all cases, and then pressed Download button as highlighted.
- The download will begin and you just have to wait.

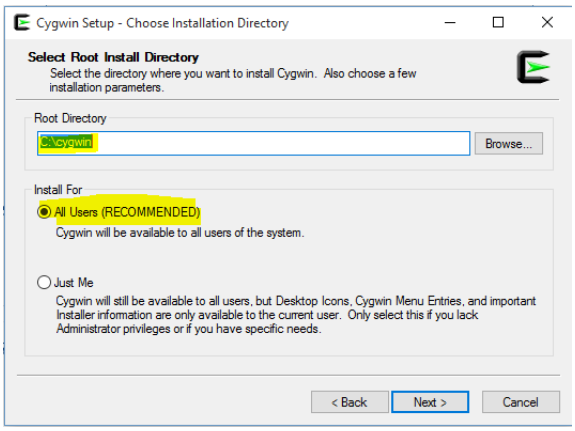
5.1 Install Cygwin 32 bit

Cygwin 32 bit version is required for this utility to work. Mind that it doesn't matter what PC type you have (32bit or 64bit). You ABSOLUTELY have to install the 32 bit version only.

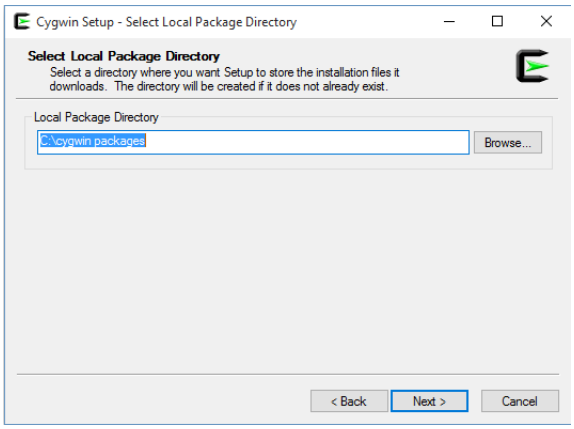

https://cygwin.com/install.html
Get that Linux feeling - on Windows
Installing and Updating Cygwin Packages
Installing and Updating Cygwin for 32-bit versions of Windows
Run setup-x86.exe any time you want to update or install a Cygwin package for 32-bit windows. The signature for setup-x86.exe can be used to verify the validity of this binary using this public key.
Installing and Updating Cygwin for 64-bit versions of Windows
Run setup-x86_64.exe any time you want to update or install a Cygwin package for 64-bit windows. The signature for setup-x86_64.exe can be used to verify the validity of this binary using this public key.
General installation notes
When installing packages for the first time, setup*.exe does not install every package. Only the minimal base packages from the Cygwin distribution are installed by default, which takes up about 100 MB.
Clicking on categories and packages in the setup*.exe package installation screen allows you to select what is installed or updated.


Cygwin Setup
Cygwin Net Release Setup Program
This setup program is used for the initial installation of the Cygwin environment as well as all subsequent updates. Make sure to remember where you saved it.
The pages that follow will guide you through the installation. Please note that Cygwin consists of a large number of packages spanning a wide variety of purposes. We only install a base set of packages by default. You can always run this program at any time in the future to add, remove, or upgrade packages as necessary.
Setup.exe version 2.873 (32 bit)
Copyright 2000-2015
http://www.cygwin.com/
< Back
Next >
Cancel

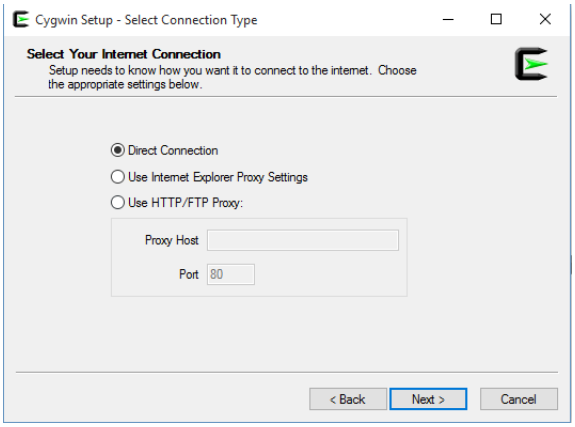

Cygwin Setup - Choose Installation Type
Choose A Download Source
Choose whether to install or download from the internet, or install from files in a local directory.
Install from Internet
(downloaded files will be kept for future re-use)
Download Without Installing
Install from Local Directory
< Back
Next >
Cancel

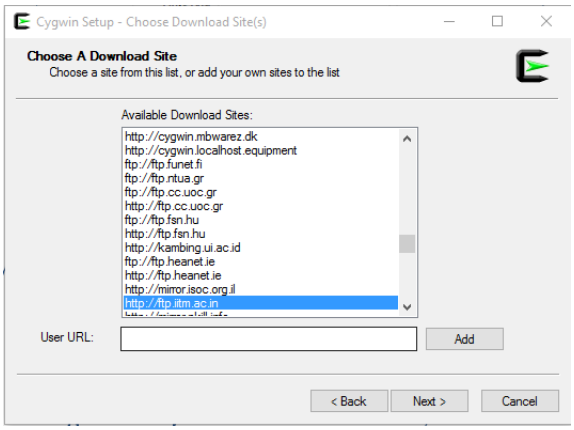


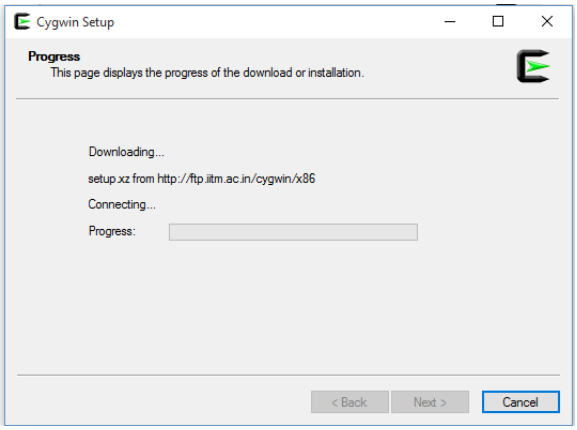
- Create a new directory Cygwin packages in C:/ and use this to store your packages, as shown below.



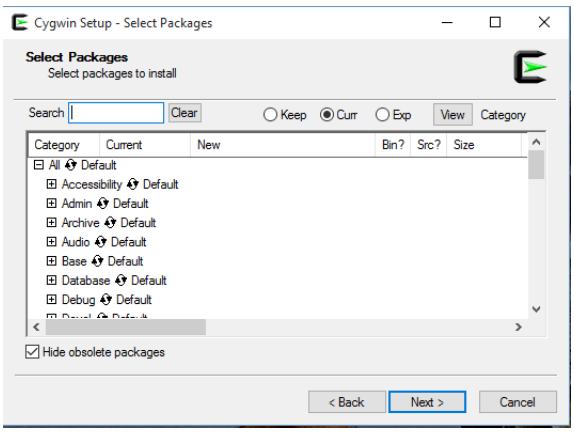
- Choose any mirror near to you.

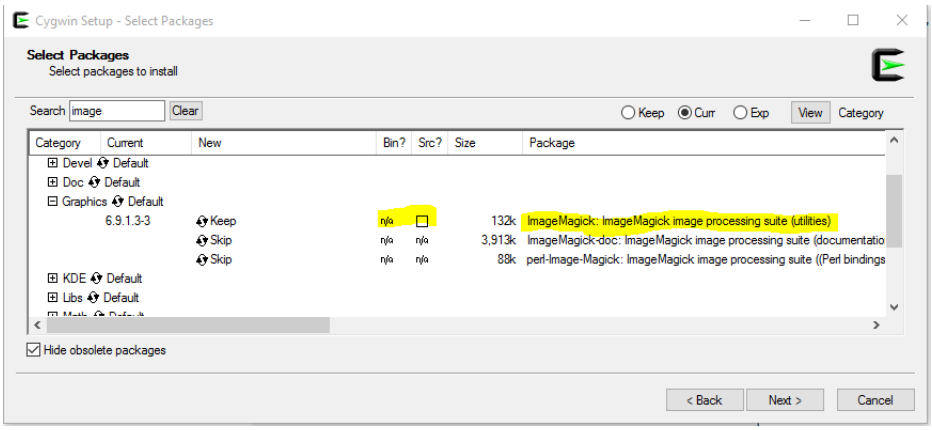




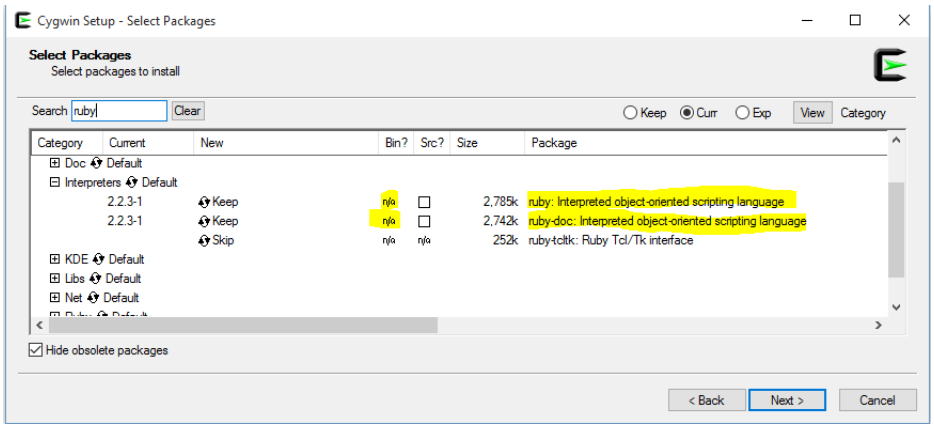


- Now we need to install two packages. Search for Image in the search bar. Then, when the results appear, go to graphics, and then click on ImageMagick. Make sure the bin checkbox is checked. src is optional.

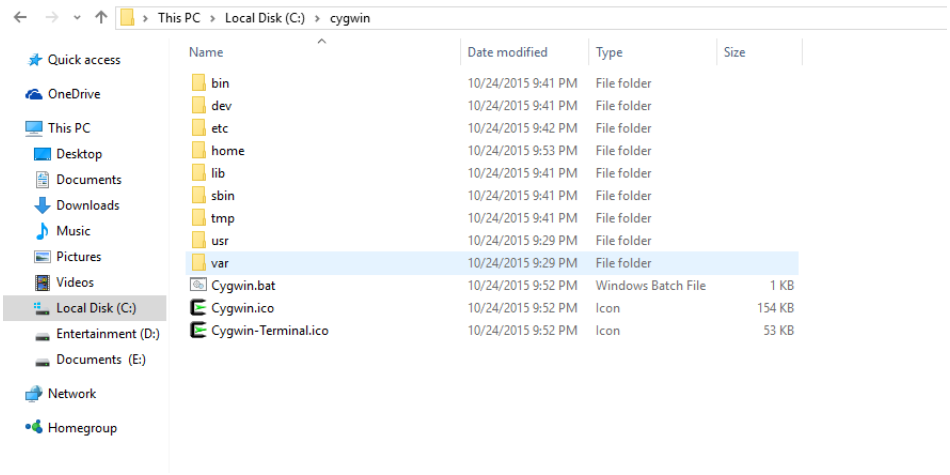




- Again search for Ruby in the search bar. Once the search results appear, go to Interpreters and select Ruby by checking its corresponding bin checkbox. These steps are shown below. My steps are different because of the fact that I already have these packages installed.



- Just click next and wait for the setup to download the required packages and install it.
- Once Cygwin is successfully installed in your computer, your C:/Cygwin/ folder will look something like this:
- Run Cygwin.bat once, and then close it once it loads.



6 Starting the Conversion Process

6.1 Confirm Cygwin works

If you've carefully followed the above procedure, you now have everything that's required for the conversion process to start. Before starting though, let's make sure you got Cygwin installed properly:

- Open cygwin.bat from the Cygwin main directory.
- just write "chmod" there and see if it prints the "help" section of the chmod command.
- Now go to bin/bash.exe and open it.
- Do step 2 here again.
- If it says "Command not found", that means Cygwin failed to update the System's PATH variable, and now you would have to do that manually.
- DO THE BELOW STEP ONLY IF YOU GOT THE ERROR ABOVE
- Put the path to your bin folder (in my case its C:/cygwin/bin/) in your PATH system variable.

6.2 Start the Process

- Go to your C:/cygwin/bin directory.
- Copy the cygwin1.dll file and paste it inside the folder where your DDMUtility is present.
- Now open the script:
  openDDSMLJPEGAndConvertToPNG.m
- In the start of the script you will see the following 3 variables:

```
%% SET THESE PATHS FIRST!

% Directory where all cases are present. Just change this directory
% folders you have, like Benign01, Benign02, ...
CollectionDirectory = 'D:\Test_For_New_Script\';

% These are for the other script. NOTICE THE BRACKETS DIRECTION. ALS
% NOTICE A SEMICOLON IN FIRST VAR.
pathToJPEGandDDSM2RAWfiles = 'D://Test_For_New_Script//;';
cygwinLocation = 'C:\\cygwin\\bin\\bash';
```

- CAREFULLY change the variables to your needs. MAKE SURE YOU PUT THE SLASHES AND DOUBLE SLASHES CORRECTLY JUST AS IT IS SHOWN. ALSO MAKE SURE ABOUT AN EXTRA SEMICOLON IN THE SECOND VARIABLE.
- DOUBLY MAKE SURE YOUR FOLDER NAME DO NOT HAVE SPACES. IT MUST BE ADJOINED USING UNDERSCORES.
- Once the variables are updated, just run the script. Your command window will look something like this:

```
Converting file A_1573_1.LEFT_CC.LJPEG...
Currently in Folder -
Benign01/case1573
Decompressing LJPEG -> LJPEG1...
Converting LJPEG1 -> RAW...
./Benign01/case1573/A_1573_1.LEFT_CC.LJPEG.1-ddsmraw2pnm.pnm
Inside case - case1573...
Saving .png file to disk...
```

- All your converted PNG files will be saved in the PNGFiles folder in each of the Case folders you will have.
- If you can see this, CONGRATULATIONS! You successfully set up the conversion procedure. Now just sit back and relax. The conversion is a system heavy process, and each image may take from 10 seconds to upto 5 minutes depending on your configuration. On my laptop, each image takes roughly 1 minute to get converted and saved to PNG format on hard disk.

7 Problems and Contact

Although the utility is very extensively tested on a number of PCs and Servers, you may run into some problems. Be advised, that most of the problems that you will face will be the direct result of negligence towards reading carefully through this tutorial. Just restart the process with a fresh mind and it will work.

If you still believe that there is some problem or there is a bug and would like to report the same, please feel free to email me at: anmol.sharma293@gmail.com