Flavia,

A Leaf Recognition Algorithm/ Program for Plant Classification using PNN (Probabilistic Neural Network)

Stephen Gang Wu (gwu827@gmail.com)
Dept. Biological Systems Engineering, Virginia Institute of Technology, USA

Forrest Sheng Bao (http://narnia.cs.ttu.edu)
Dept. of Computer Science/Dept. of Electrical Engineering, Texas Tech University, USA

Eric You Xu (http://youxu.info)
Dept. of Computer Science and Engineering, Washington University in St. Louis, USA

Special thanks to:
Lavi Shi Chen (http://www.ento.psu.edu/Personnel/students/shi_chen.html)
Dept. of Entomology, School of Agriculture, Pennsylvania State University, USA

Please cite this paper if you use our data and program in your publication: Stephen Gang Wu, Forrest Sheng Bao, Eric You Xu, Yu-Xuan Wang, Yi-Fan Chang and Qiao-Liang Xiang, A Leaf Recognition Algorithm for Plant classification Using Probabilistic Neural Network, IEEE 7th International Symposium on Signal Processing and Information Technology, 2007, Cairo, Egypt.

Last modified: Dec. 24, 2009 Forrest Sheng Bao

This research was conducted under the supervision of

Prof. Xingjun Tian (tianxj@nju.edu.cn), S. of Life Science, Nanjing University

This research was supported by

Dept. of Botany, School of Life Science, Nanjing University

1. Introduction

The purpose of this program is to teach a computer to classify plants via their leaves. You just need to input the leaf image of plant (acquired via digital camera or scanners), then the computer can tell you what kind of plant it is.

We utilize the PNN (Probabilistic Neural Network) to implement this AI process. 12 characters of leaves are taken into account, including geometrical ones and morphological ones. After discriminant analysis (stepwise method), all these characters are reserved. PCA orthogonalizes these 12 characters into 5 principal variables, which are input vectors of the PNN.

Presently, our system can classify 32 common plants in Yangtze Delta(Shanghai is just there), China, such as (in Latin name) *Phyllestachys Pubescens*, *Aesculus Chinensis*, *Berberis Ferdinandi-coburgii Schneid*, *Cercis Chinensis*, *Indigofera Tinctoria Linn.*, *Acer Dalmatum*, *Phoebe Zhennan*, *Kalopanax Septemlobus*, *Cinnamomum Japonicum*, *Koelreuteria Paniculata*, *Ilex Macrocarpa*, *Pittosporum Tobira*, *Chimonanthus Praecox*, *Cinnamomum camphora*, *Viburnum Awabuki*, *Osmanthus Fragrans*, *Cedrus Deodara*, *Ginkgo Biloba*, *Lagerstroemia Indica*, *Nerium Indcus*, *Podocarpus Macrophyllus*, *Prunus Yedoensis*, *Ligustrum Lucidum Ait.*, *Toona Sinensis*, *Amygdalus Persica*, *Manglietia Fordiana*, *Acer Buergerianum*, *Mahonia Bealei*, *Magnolia Grandiflora*, *Populus Xcanadensis Moench*, *Liriodendron Chinense*, *Citrus Reticulata Blanco*. The average accuracy is 93% for all of them.

Homepage of this program: http://flavia.sourceforge.net/

2. Requirements to run this program

- MATLAB R14(7.0) or later.
- MATLAB Image Processing Toolbox and Neural Network Toolbox.
- Java Virtual Machine to enable the graphic user interface of MATLAB.

* We have only tested *flavia* on MATLAB for Linux (OpenSUSE 10.2, Mandriva 2007, Ubuntu 6.10/7.04, CentOS 5/Fedora 6). So we are not sure whether it could work well on other operating systems, such as Windows. Here is the version information of MATLAB that we used in this manual. The Linux distribution is Ubuntu Linux 7.04 (codename “Feisty Fawn”).

```
Operating System: Linux 2.6.20-15-generic #2 SMP Sun Apr 15 07:36:31 UTC 2007 i686
Java VM Version: Java 1.5.0 with Sun Microsystems Inc. Java HotSpot(TM) Client VM mixed mode
--------------------------------------------------------------------------------
MATLAB                                                  Version 7.4        (R2007a)
Image Processing Toolbox                                Version 5.4        (R2007a)
MATLAB Builder for Java                                 Version 1.1        (R2007a)
MATLAB Compiler                                         Version 4.6        (R2007a)
Neural Network Toolbox                                  Version 5.0.2      (R2007a)
Signal Processing Toolbox                               Version 6.7        (R2007a)
```

To run in Linux binary executables, you will need these dynamical libraries: `linux-gate.so.1`, `libmwmclmcrrt.so.7.5`, `libm.so.6`, `libstdc++.so.6`, `libpthread.so.0`, `libc.so.6`, `ld-linux.so.2`, `libgcc_s.so.1`. The `libmwmclmcrrt.so.7.5` is provided by MCR which will be mention

later. Other libraries are general Linux libraries, provided by most up-to-date Linux distributions.

There are some tabs embraced with a pair of "<>", like <*version*>, <*arch*> or <*OS*>. Please replace them with proper strings matching your computer, such as <*version*>=alpha, <*arch*>=x86_64, <*OS*>=Linux.

3. File structure and distribution

We distribute *flavia* in two parts: the main program (in form of MATLAB script or Linux binary executables) and the standard JPEG images of plants that can be classified by *flavia*. Complete raw dataset of 1GB leaf image are also available for downloading on project website.

3.1 The main program

This main program is distributed as MATLAB files. You can consider the MATLAB scripts as "source code." Please download `flavia-<version>.src.tar.gz` package if you want the MATLAB scripts. Otherwise you can download the binary package.

3.2 The Standard JPEG Images

Beside Linux binary executables and MATLAB scripts, we also distribute JPEG images of plants that can be classified by this program. They are acquired by our own digital camera in our labs. The JPEG images of standard leaves are used for a side-by-side display comparison between the image you inputed and the image of standard leaf. This is an optional function of *flavia*. It provides users a way to determine whether the computer has recognized the plant correctly. If it recognized the plant correctly, you will see that the leaf image you entered is the "same" as the image of standard leaf, as Fig. 1 shows. If you don't need this function, just download the main program. If you need it, remember the name is like "`standardleaves.tar.bz2`"

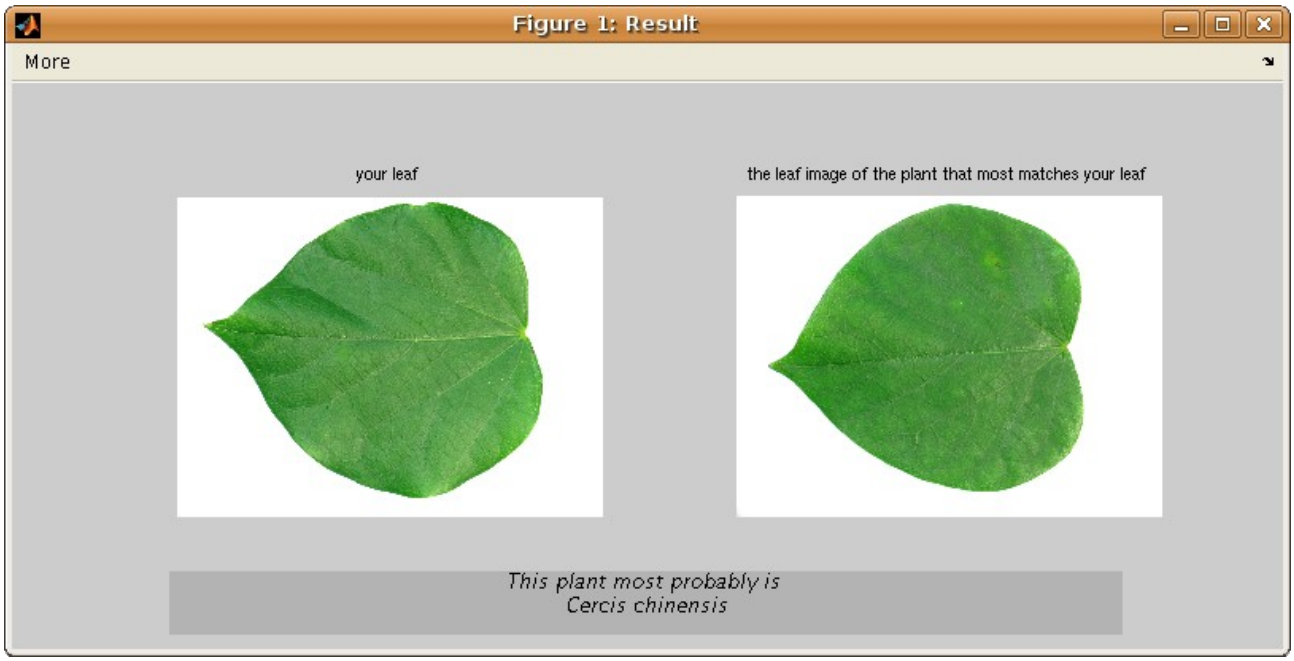

Figure 1: Side-by-Side display between your image and standard leaf

The standard leaf images are supposed to locate in a folder called `standardleaves` which is at the same level directory with MATLAB scripts or binary executables of *flavia*. The *flavia* program will search this folder from the place where it locates. That directory contains many JPEF files that has been numbered. We won't mention where to put it later in this document. So keep in mind here.

4. Downloading and preparing files (source, MATLAB scripts)

Here we use "$FLAVIA" to denote the directory that you put all M-files and MAT-files. It is just the directory where you want to run *flavia*. In following examples, $FLAVIA is just `/forrest/work/cs/image` .
To run *flavia* on MATLAB environment(as MATLAB scripts), you just need to extract M-files(MATLAB script) and other required files(MAT file, JPEG images, etc.) to $FLAVIA. We use the term "MATLAB files" to refer both MATLAB scripts and MAT files which store data. To *flavia*, all MATLAB files are located on the same hierarchy level. The directory, $FLAVIA, should contain following files:

```
forrest@flavia:/forrest/work/cs/image$ ls
DetailInfo.m      imagemd.m       SoftInfo.m
five.mat          input.mat       taxonomynocn.mat
getpts1.m         net2.mat        pcapnn.m
flavia.m          seg2.mat
```

If you want to enable the side-by-side display comparison function, you should also see a folder called "`standardleaves`". You can refer to Sec. 3.2 for details about it.

When required files are ready, you can start MATLAB and enter that directory in MATLAB prompt. You can also use the GUIs provided by MATLAB to enter that directory. After entering $FLAVIA, just type "flavia" and press "ENTER" key. For example, on my computer, I just run(don't type the ">>"):

```
>> cd /forrest/work/cs/image
>> flavia
```

Here is a snapshot illustrating how to start flavia as MATLAB scripts. You can see commands I entered on the command windows. On the left banner you can see some required files to run flavia. (Some extra files and folders are not needed actually. They are just my working things.) If you could see that pop-up dialog prompting you to input the leaf image, it means OK. Congratulations, you can now move to Sec. 5.

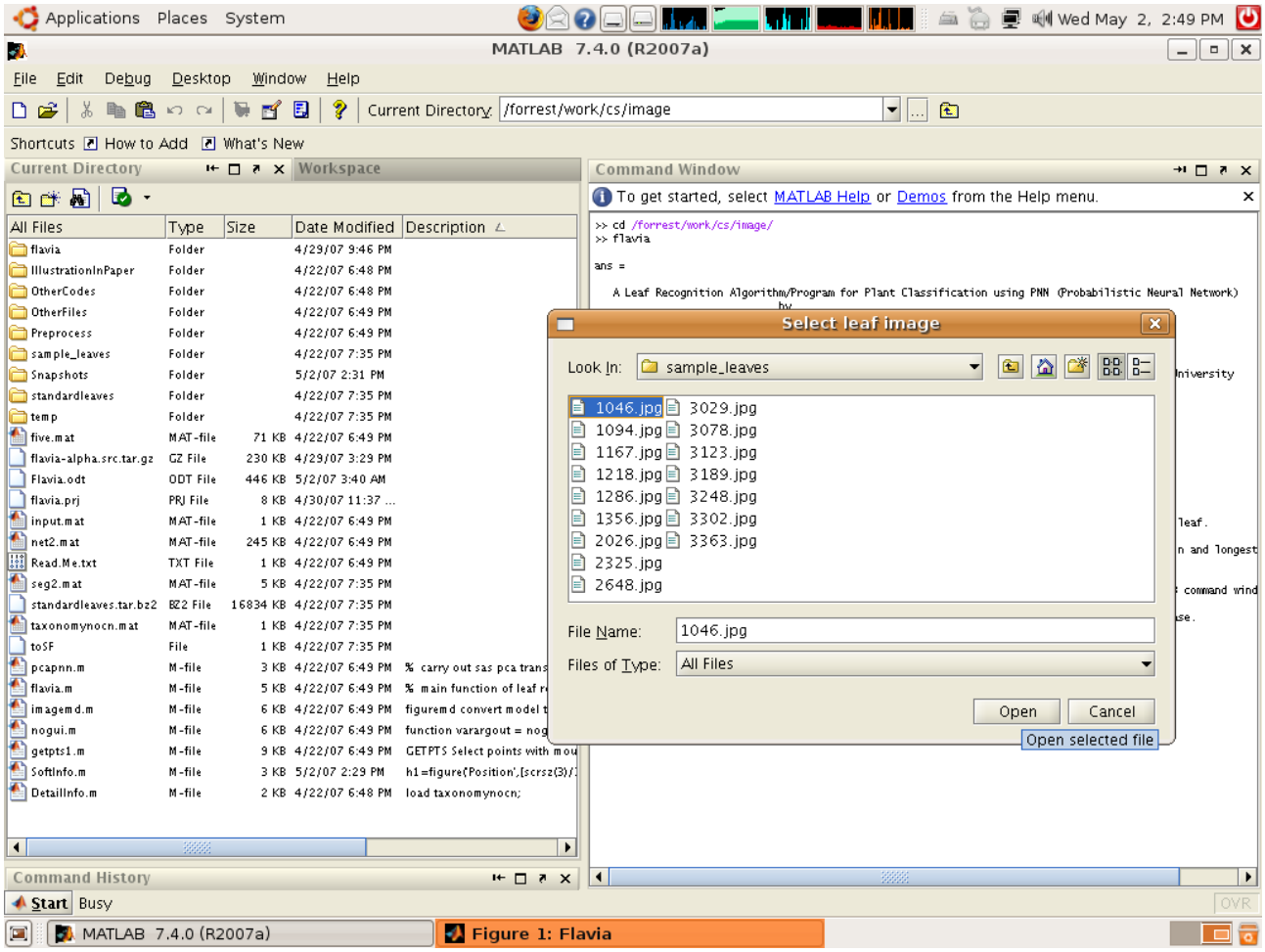

Figure 2: Starting Flavia as MATLAB scripts

5. Step-by-step to run *Flavia*

Here we take MATLAB scripts version of *flavia* as an example.

After starting *Flavia* as Fig. 2, you can select and open the leaf image you want to classify. You should also see a newly-opened window providing basic information about *flavia*, as Fig. 3. You can click “More” ->”Details about our algorithms” or click the “Details of this algorithm” button for details as show in Fig. 4

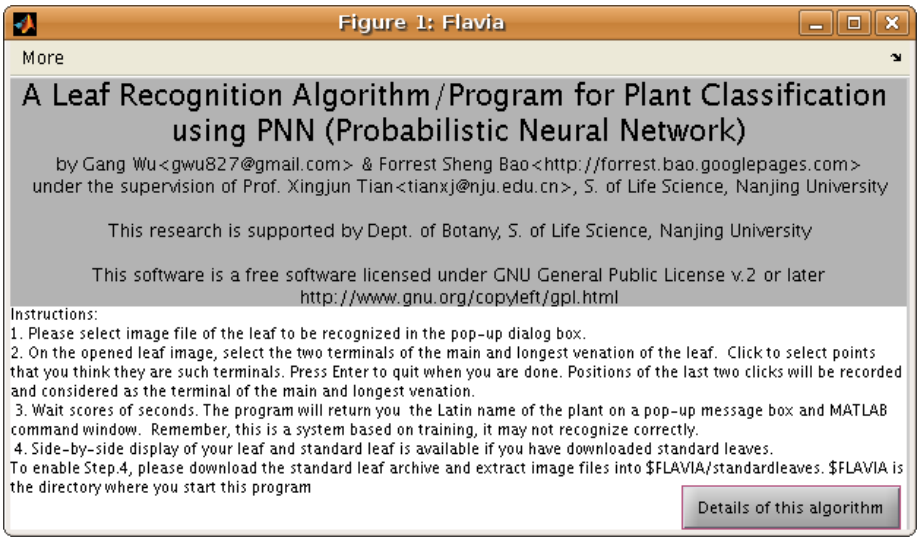

Figure 3: the window showing information about Flavia

Attention: We have found some display problem due to fonts and display resolution. If you can't see some windows completely, please report to us about your font type, size and your screen resolution.

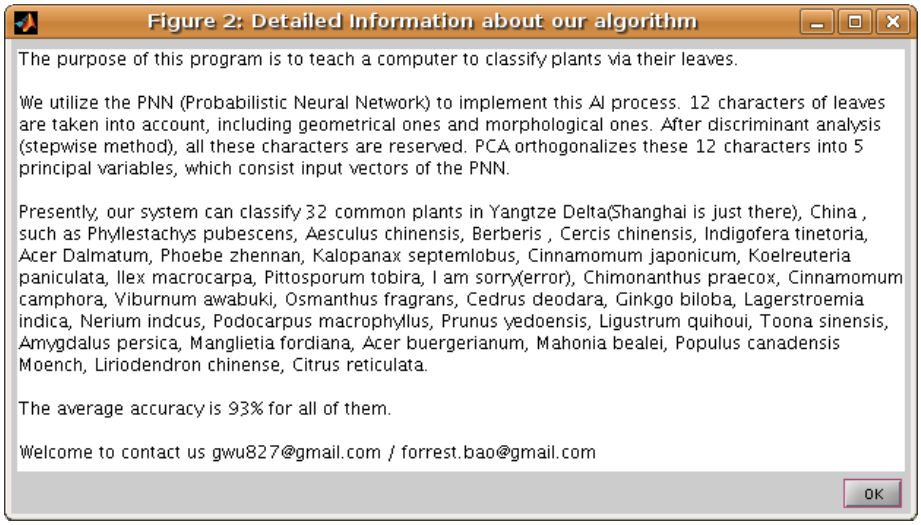

Figure 4: the window showing details about our algorithm

The only “non-automatic” part of our algorithm is that you need to mark the two terminals of the longest or the main venation of the leaf. After you have selected an image to be processed, a new window containing the image of your leaf will pop-up to allow you select these two terminals. The mouse pointer becomes like this :

Click the primary key(left button to right-handed mouse) of your mouse to select. There will be an asterisk( ) in the place where you click. If you make wrong click, don't worry. Only the last two selections will be recorded and considered as the two terminals. Fig. 5 is an example of selecting those two terminals. You can see two asterisk on the snapshot. When you are done, press ENTER key on your keyboard. Now everything

else will be processed by the program. You are free.

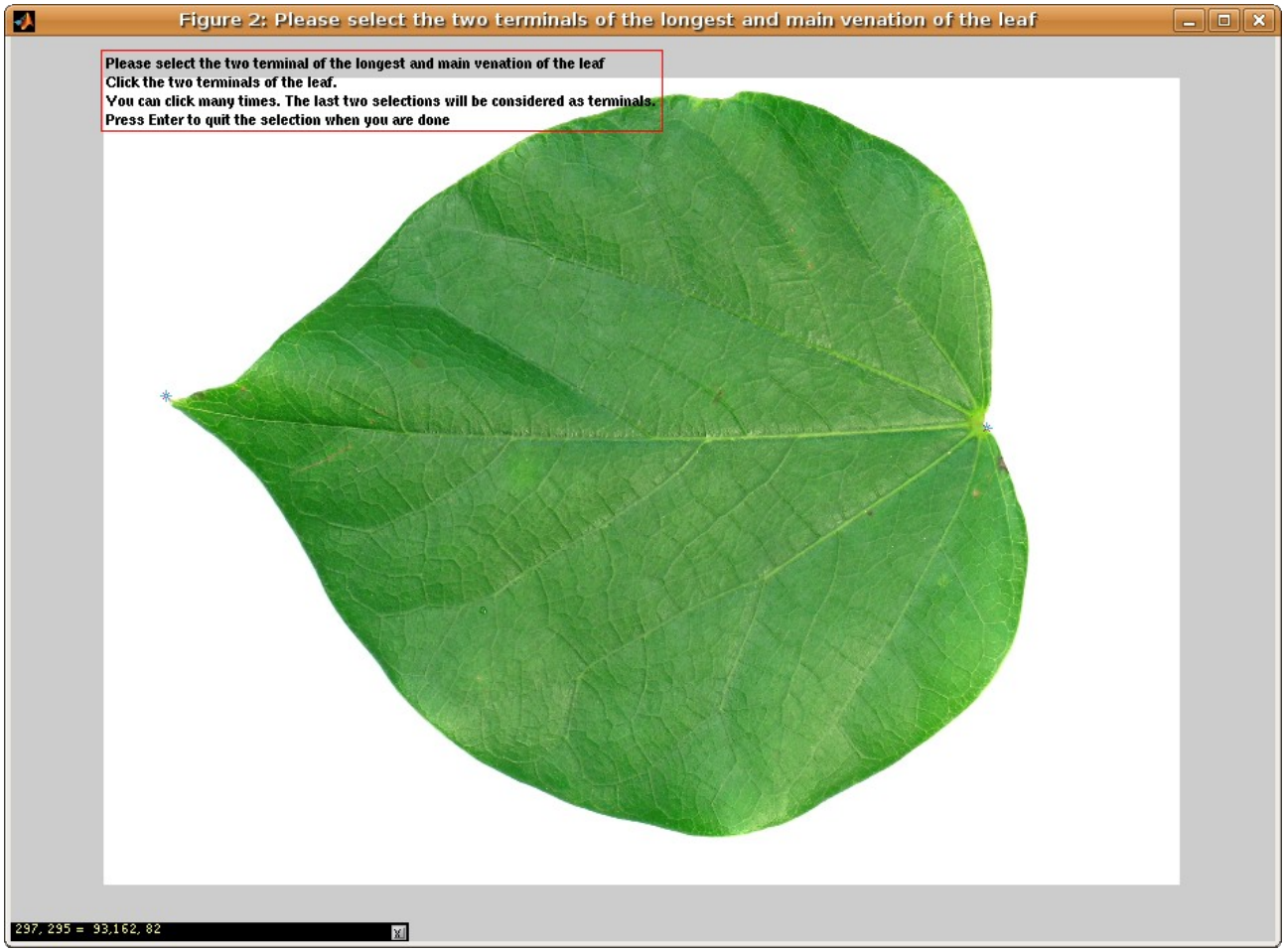

Figure 5: Click to select the terminals

When the program is done, you can see a pop-up window telling you what kind of plant it is, as Fig. 6. The name is not in English but Latin. If you have enabled the side-by-side display option(by putting standard leaf image into `standardleaves` folder, as mentioned in Sec. 2.2), you will also see another window as Fig.1

Figure 6: Result box

More detailed log will be displayed on the MATLAB command window, like this:

```
ans =
              recognition started at 04 01, 2007 16:10:11.320
classification_result =
                                        4
Elapsed time is 0.377743 seconds.
ans =
             recognition finished at 04 01, 2007 16:10:11.702
msgresult =
            the Latin name of this plant is: Cercis Chinensis
```

Attention: If it returns the plant name is “I am sorry” then it means that our program can't classify the plant

you inputed.

6. Last but not least...

Flavia is still an alpha version software. Both the program and this document are not completed. More details about our algorithm, such as parameters will be discussed later. You are warmly welcomed to give us any suggestion. Please feel free to leave lines to us. Since we are all very busy on our own research and study, we planned to pause this project for a while. So now we don't have any plan to enlarge the number of plants we can classify. We don't wanna modify the structure of neural networks either.

If you need help, please refer to our homepage for help, e.g., mailing list or forum.

Again, please cite our paper, if you use our data or program in your publication: Stephen Gang Wu, Forrest Sheng Bao, Eric You Xu, Yu-Xuan Wang, Yi-Fan Chang and Qiao-Liang Xiang, A Leaf Recognition Algorithm for Plant classification Using Probabilistic Neural Network, IEEE 7th International Symposium on Signal Processing and Information Technology, 2007, Cairo, Egypt.