

MRC, HARWELL

---

Systems Imaging Group

# HARP User Manual

# Table of Contents

Introduction .....	1
Getting Started .....	2
Windows .....	2
Linux .....	2
How to use HARP .....	3
Overview .....	3
Input and output select.....	4
Imaging Modality Selection .....	5
Recon File Info.....	5
Identification.....	5
Additional files and folders .....	7
Options.....	7
OPT channels .....	11
Status section .....	11
Add to Processing List .....	12
Processing Tab .....	12
Output Summary .....	12
Appendix .....	14
7-Zip instructions.....	14
Required external python modules .....	15

## Introduction

This manual will give a step by step guide on how to use HARP (Harwell Automated Recon Processor). The tool has been developed by the Systems Imaging Group at MRC Harwell as an aid in the processing of  $\mu$ CT and OPT scans and reconstructions.

The intended end user is an employee working on the IMPC embryo phenotyping pipeline.

For any queries please contact any of the following:

Henrik Westerberg: [h.westerberg@har.mrc.ac.uk](mailto:h.westerberg@har.mrc.ac.uk)

Neil.Horner: [n.horner@har.mrc.ac.uk](mailto:n.horner@har.mrc.ac.uk)

Tom Lawson: [t.lawson@har.mrc.ac.uk](mailto:t.lawson@har.mrc.ac.uk)

## Getting Started

### Windows

HARP has been developed to work in either Linux or Windows environments. Below describes the steps to get HARP up and running.

**To run HARP from a Windows executable** simply copy the latest version of HARP into the desired location. Then navigate to the Main.exe file and double click on the file. Alternatively a shortcut can be made by right clicking on the Main.exe file, selecting “createshortcut”. This shortcut can be copied to a convenient location.

**Box 1: HARP executable location**

```
C:\soft\HARP\dist\Main.exe
```

**To run HARP directory from Python script** a number of python modules will need to be installed. See Appendix for full list. Once the modules are installed, copy the HARP scripts to the desired location and run the Main.py in python

### Linux

HARP can currently be run as a python script in Linux. The dependent modules need to be installed. Once the modules are installed, copy the HARP scripts into the desired location and run the Main.py in python.

**Box 1: Run HARP as python Script**

```
C:\soft\HARP\dist\Main.exe
```

## How to use HARP

This section will take you through each step of how to process your data with HARP. The screenshots are taken from the Windows version of the software so it will look different if HARP is running in Linux.

### Overview

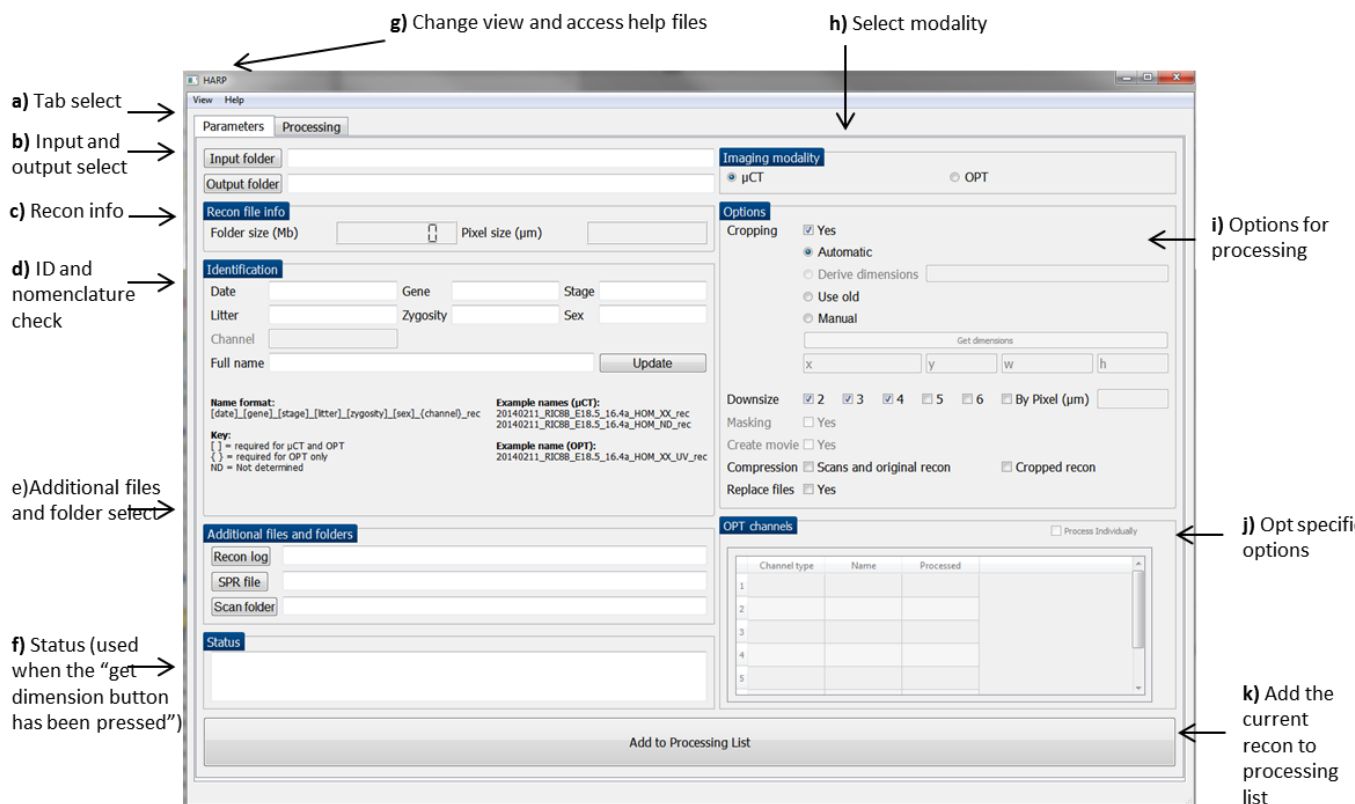


FIGURE 1: Starting view of HARP. The current tab is selected on "Parameters". Any functionality not currently available is greyed out.

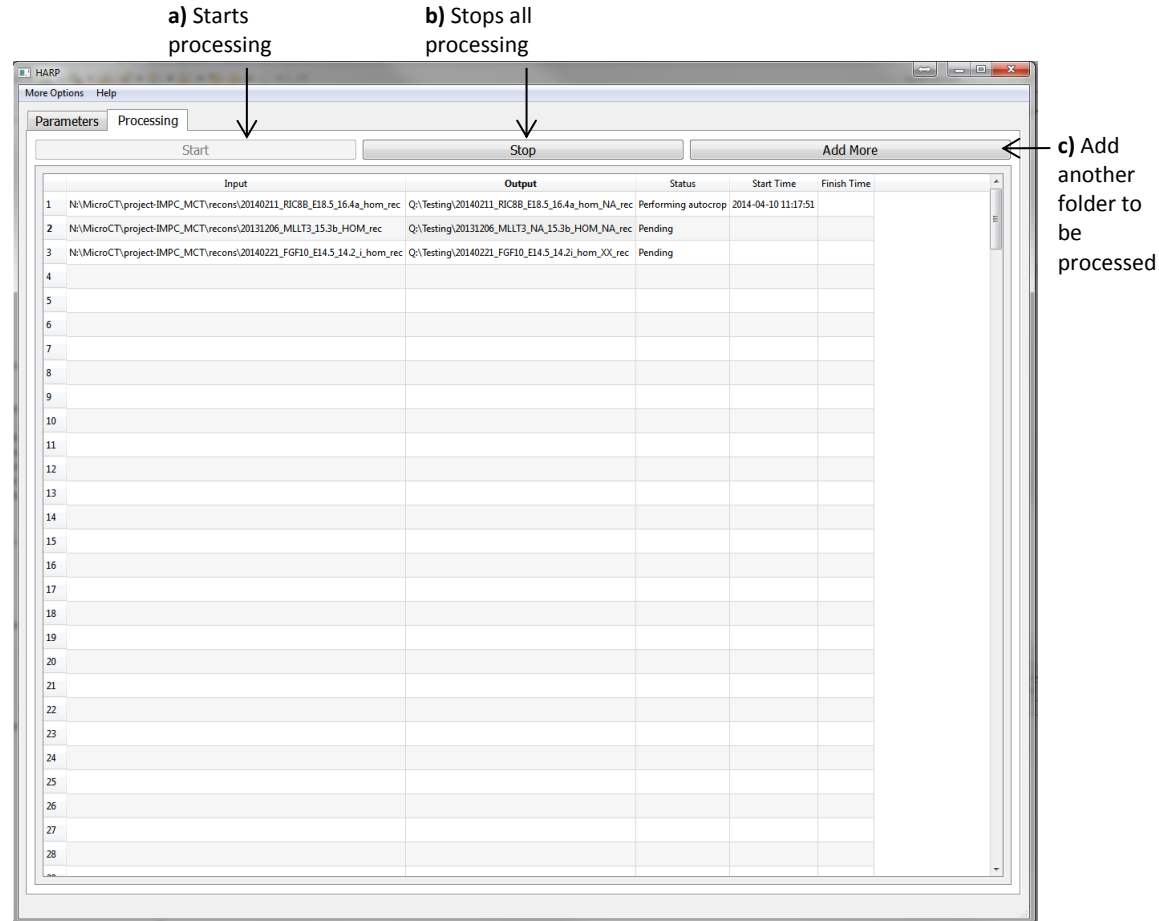


FIGURE 2: Processing view of HARP, shown when the “Processing” tab is selected and whenever as user successfully adds a recon folder to the processing list.

## Input and output select

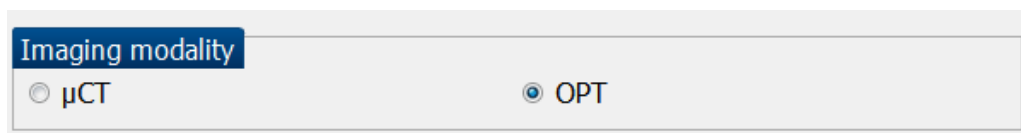
The first step in any use of the HARP will be to select the files to be processed. This is achieved by selecting the input folder button or drag and dropping a recon folder into HARP. The user should select the input directory which contains all reconstructed images which are required for processing.

After a folder have been selected HARP attempts to auto-populate a number of fields. This includes suggesting an output folder, identifying if OPT or  $\mu$ CT, calculating the recon folder size, identifying the pixel size, checking the identification nomenclature, locating the recon log fie, locating the SPR file and locating the relevant scan folder.

If the folder being processed is from the IMPC pipeline, the output folder is automatically assigned to the “processed recons” directory in the project-IMPC directory. Alternatively the output folder can be assigned manually by clicking on the “output folder” button

## Imaging Modality Selection

HARP can use  $\mu$ CT or OPT data. HARP will try and identify this automatically but it can be manually selected.



The interface shows a tab labeled "Imaging modality". Below the tab are two radio buttons. The first is labeled " $\mu$ CT" and is unselected. The second is labeled "OPT" and is selected, indicated by a blue dot.

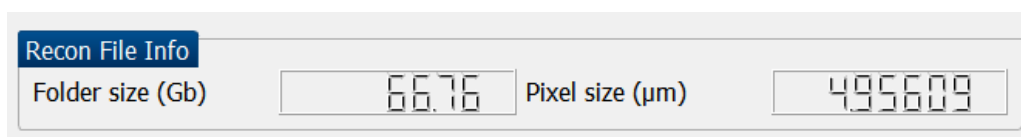
FIGURE 4: Imaging modality selection.

## Recon File Info

The recon file info is updated when an input folder is selected.

This section lets the user know what the folder size is of the original recon directory. The folder size calculation is based on only the recon files and does not include any subdirectories.

The pixel size is determined from the associated recon log file.



The interface shows a tab labeled "Recon File Info". Below the tab are two input fields. The first is labeled "Folder size (Gb)" and contains the value "66.76". The second is labeled "Pixel size ( $\mu$ m)" and contains the value "495609".

FIGURE 5: Recon file information.

## Identification

The identification section allows the user to see if the nomenclature of the reconstruction is in a standard format. The format includes the information date, gene, stage, litter, zygosity and sex separated by an underscore. If one of the identifiers is not known the letters NA or ND should take its place. OPT naming should also include the channel. See below for naming conventions.

### Box 3: $\mu$ CT Nomenclature

#### Name format:

[date]\_[gene]\_[stage]\_[litter]\_[zygosity]\_[sex]\_rec

#### Example names:

20140211\_RIC8B\_E18.5\_16.4a\_hom\_XX\_rec

20140211\_RIC8B\_E18.5\_16.4a\_hom\_ND\_rec

#### Box 4: OPT Nomenclature

##### Name format:

[date]\_[gene]\_[stage]\_[litter]\_[zygosity]\_[sex]\_[channel]\_rec

##### Example names:

20140211\_RIC8B\_E18.5\_16.4a\_hom\_XX\_UV\_rec

20140211\_RIC8B\_E18.5\_16.4a\_hom\_XX\_W\_rec

If the wrong name format has been used a warning box will pop up. The user can carry on with this naming format or can change the name for all processed data.

**Identification**

Date	20140514	Gene	TULP3	Stage	E18.5
Litter	16.4a	Zygosity	MUTANT	Sex	XX
Channel	W				

Full name 20140514\_TULP3\_E18.5\_16.4a\_MUTANT\_XX\_W\_rec

**Name format:**  
[date]\_[gene]\_[stage]\_[litter]\_[zygosity]\_[sex]\_{channel}\_rec

**Key:**  
[ ] = required for μCT and OPT  
{ } = required for OPT only  
ND = Not determined

**Example names (μCT):**  
20140211\_RIC8B\_E18.5\_16.4a\_HOM\_XX\_rec  
20140211\_RIC8B\_E18.5\_16.4a\_HOM\_ND\_rec

**Example name (OPT):**  
20140211\_RIC8B\_E18.5\_16.4a\_HOM\_XX\_UV\_rec

Where to edit  
name

Then update

**Identification**

Date	20140514	Gene	TULP3	Stage	E18.5
Litter	16.4a	Zygosity	MUT	Sex	XX
Channel	W				

Full name 20140514\_TULP3\_E18.5\_16.4a\_MUT\_XX\_W\_rec

**Name format:**  
[date]\_[gene]\_[stage]\_[litter]\_[zygosity]\_[sex]\_{channel}\_rec

**Key:**  
[ ] = required for μCT and OPT  
{ } = required for OPT only  
ND = Not determined

**Example names (μCT):**  
20140211\_RIC8B\_E18.5\_16.4a\_HOM\_XX\_rec  
20140211\_RIC8B\_E18.5\_16.4a\_HOM\_ND\_rec

**Example name (OPT):**  
20140211\_RIC8B\_E18.5\_16.4a\_HOM\_XX\_UV\_rec

FIGURE 5: Identification section. A orange circle has been added for this image to highlight where a non-canonical name has been used. A blue circle identifies where the name has been edited and the black circle identifies the change which is made after “update” is pressed. The naming of all future files and folders is also updated when the “update” button is pressed.

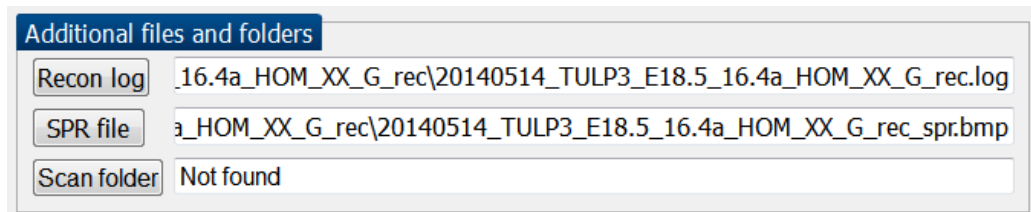


## Additional files and folders

HARP automatically identifies the recon log, SPR file and the scan folder. The recon log file is required if the recon is to be downsized by a pixel.

The scan folder is required if the user wants to compress the scan folder.

If a folder can't be automatically located the text "Not found" will be shown. Processing can still continue with this message.



Additional files and folders

Recon log: \_16.4a\_HOM\_XX\_G\_rec\20140514\_TULP3\_E18.5\_16.4a\_HOM\_XX\_G\_rec.log

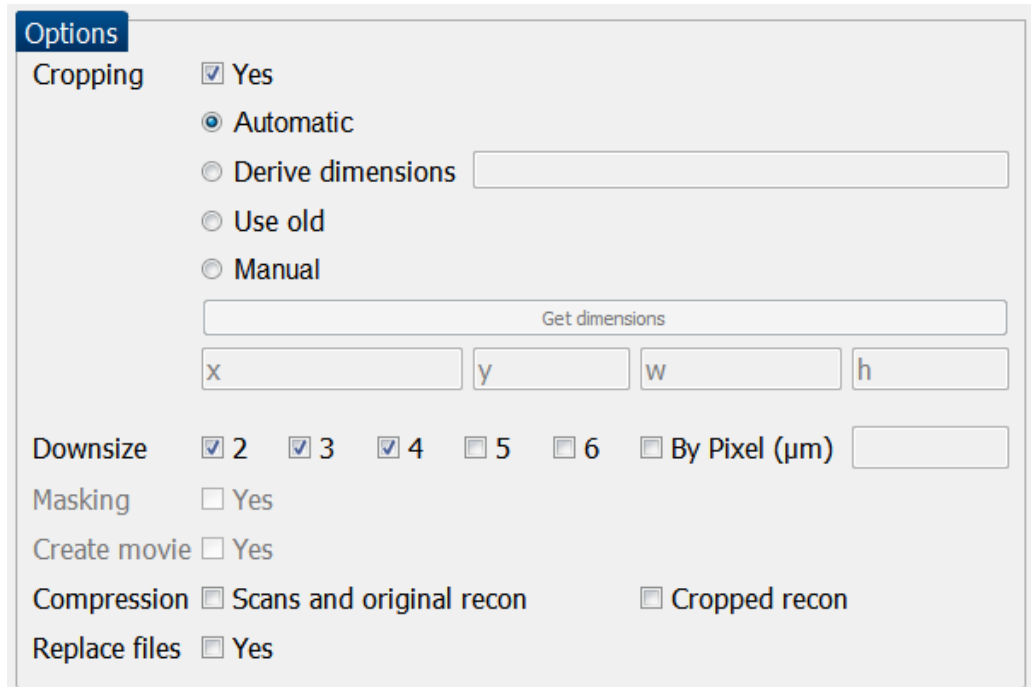
SPR file: \_a\_HOM\_XX\_G\_rec\20140514\_TULP3\_E18.5\_16.4a\_HOM\_XX\_G\_rec\_spr.bmp

Scan folder: Not found

FIGURE 7: Additional files and folders section.

## Options

User can decide on cropping, downsizing and compression options.



Options

Cropping ☒ Yes ☒ Automatic ☐ Derive dimensions ☐ Use old ☐ Manual

Get dimensions

x y w h

Downsize ☒ 2 ☒ 3 ☒ 4 ☐ 5 ☐ 6 ☐ By Pixel (µm)

Masking ☐ Yes

Create movie ☐ Yes

Compression ☐ Scans and original recon ☐ Cropped recon

Replace files ☐ Yes

FIGURE 8: Options section.

The **cropping** of images can be deselected by un-ticking the “Yes” box. When this box is not ticked HARP will downsize the original reconstruction image.

When the “Yes” box is ticked the following options are available for cropping

- **Automatic:** The cropping dimensions or crop-box will be determined automatically in the X and Y dimensions. The Z dimension will not be affected.

**Derive dimensions (OPT only):** This option provides additional functionality that for most instances is not required.

The option allows the crop dimensions to be derived from a previously processed channel. In the OPT channels section the user should select "Return" on the keyboard on the appropriate channel.

- **Use old:** If HARP has previously been used to perform the cropping of a recon the cropped image can be used to perform additional processes such as downsizing.

e.g. The user has already cropped and downsized by 3 and 4 but later realized they needed to downsize by 5 and 6 as well. The user would set up HARP as before but also select downsize by 5 and 6 and select “Use old”. This way the cropping step is not repeated.

- **Manual:** In cases where the automatic crop is not capable of determining the crop dimensions the user can they can either enter the dimensions into the x, y, width and height text boxes or press the **“Get dimensions”** button.

The **“Get dimensions”** button provides a Z-projected image of the stack, brightest point through the stack, and allows the user to select a cropping region.

The **downsizing** of the image is carried out in ImageJ and is performed after the cropping. The downsizing can be performed by a factor of 2 to 6 or alternatively a desired pixel size can be specified for the downsized image. No interpolation is used but the images are averaged when downscaling by a factor between 2 to 6. When downscaling by a desired pixel size, Bicubic interpolation is used. An individual tiff stack is created for each downsized selection.

The **compression** options allow the user to archive and compress the scan, original recon, and cropped folder. The archiving is carried out by “tar” and the compression is carried out with “Bzip”. If a Windows system is being used we recommend using **7-Zip** to access the archived and compressed folders. It should be noted here that the folder being compressed will **not** be deleted. See Appendix for 7-Zip instructions.

When the **“replace files”** option is checked the warning flags suppressed when the output folder already exists. The result of this means that items in the output folder might be replaced if they have the same name as the new files being created.

The **“create movie”** and **“mask”** options are not currently available.

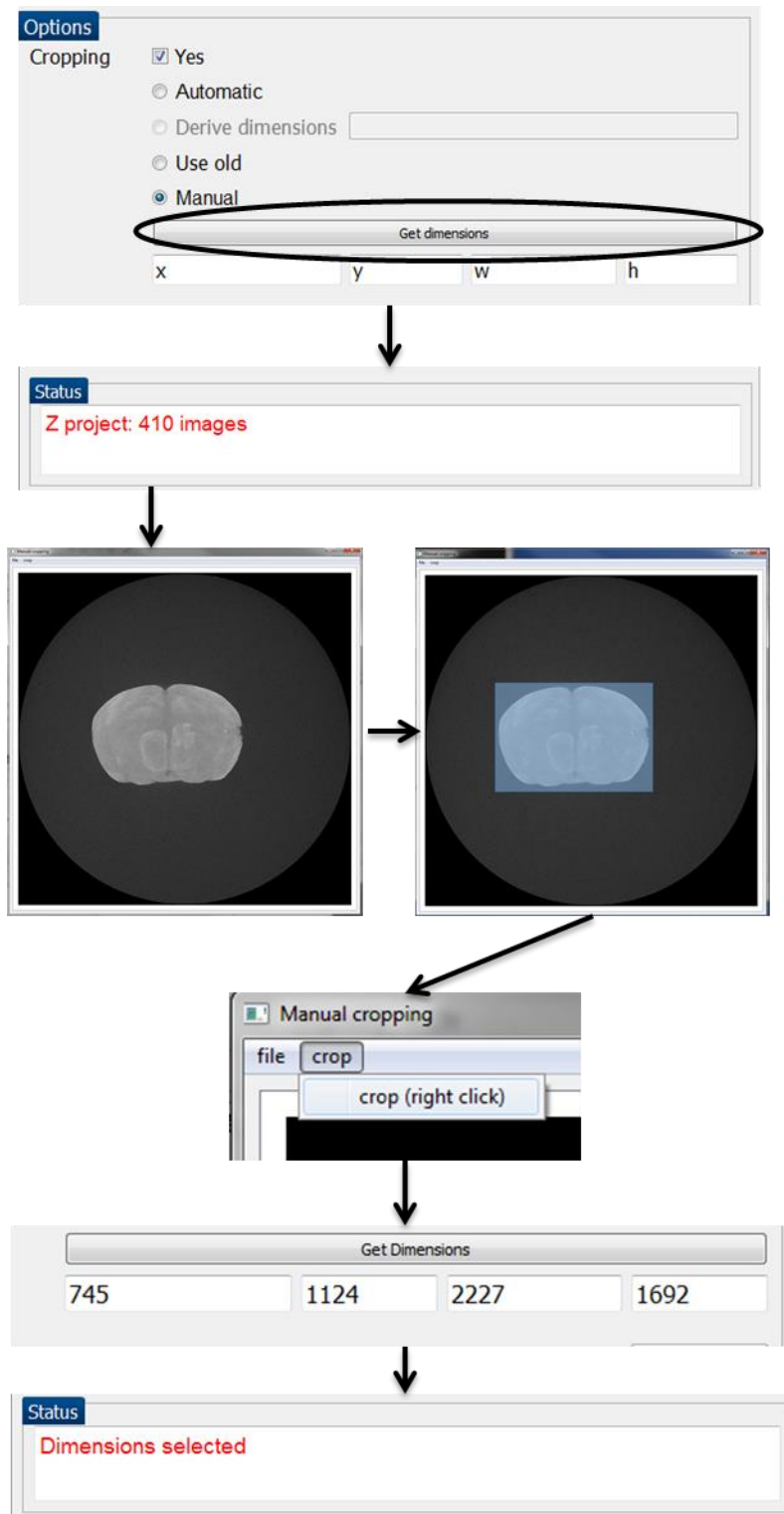


FIGURE 9: “Get dimensions” Steps to get dimensions manually from a Z-projection image.

# OPT channels

OPT channels

☐ Process Individually

	Channel type	Name	Processed
1	W	20140514_TULP3_E18.5_16.4a_HOM_XX_W_rec	No
2	UV	20140514_TULP3_E18.5_16.4a_HOM_XX_UV_rec	No
3	R	20140514_TULP3_E18.5_16.4a_HOM_XX_R_rec	No
4	G	20140514_TULP3_E18.5_16.4a_HOM_XX_G_rec	No
5			

FIGURE 9: OPT channels section.

This section displays all the OPT channels HARP could identify for the selected recon folder. If the folder is not in the correct naming format the OPT channels will not be displayed here. By default all OPT channels will be added to the processing list together and use the channel which is highlighted in red to determine the crop box dimensions.

For the example in the figure there are 4 channels: UV, White, red and green light. The white light is highlighted as red. This means the crop box dimensions will be determined using the white light channel. All other channels will derive their cropping dimensions from this channel. The user can double click on any channel to change which channel should be used to determine the crop box dimensions though.

Alternatively if this **"Process Individual Box"** is pressed the channels can be processed individually. The user can then choose the cropping method used for each channel.

The **"processed column"** details whether the channel is either already processed or on the processing list.

# Status section

This section just updates the user about what HARP is doing when the "Get Dimensions" button is clicked.

Status

Dimensions selected

FIGURE 10: Status section.

## Add to Processing List

When all the parameters have been assigned the **“Add to Processing List”** should be clicked. This will generate a folder at the specified location for the processing to be performed. Multiple folders of reconstructed images can be lined up to be processed.

The same output folder cannot be used more than once in the processing list.

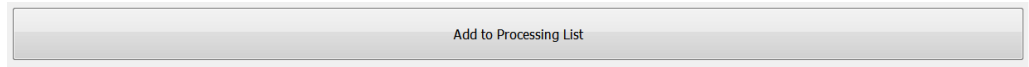


FIGURE 11: Add to Processing List button

## Processing Tab

The **“Processing”** tab displays a list of the recons which are to be processed. At the top of the tab there are three options **“Start”**, **“Stop”** and **“Add More”**. Below these buttons there is a table which displays the recons to be processed. See FIGURE 2.

When the **“Start”** button is clicked processing will begin starting with the first recon in the table which has a status of **“Pending”**. The **“Start”** button will then become greyed out until the processing has finished.

The **“Stop”** button can be clicked at any stage following the start of processing and will stop any processing currently being performed and stop further processing. Any files which that were generated prior to pressing **“stop”** will still be saved.

The **“Add more”** button should be pressed when an additional recon folder is required to be processed. It simply opens up the **“Parameters”** tab.

The Processing list table consists of 5 Columns. The **“Input”** column displays the location of the original recon. The **“Output”** column consists of the output folder. The **“Status”** column lets the user know what processing is being performed. The **“Start Time”** and **“Finish Time”** record how long the processing took.

To delete or remove a row the user should click on the row to be removed and the **“delete”** key on the keyboard should be pressed. This will remove the selected recon folder from being processed. If the recon folder has already been processed this will just remove the information from the table but will not remove any files that have been created.

## Output Summary

The output from HARP will be stored in the specified output folder. It will consist of a **“Metadata”**, **“Cropped”** and **“Scaled stacks”** sub directory (unless no cropping or scaling was performed). If the **“compress cropped recon”** option was selected then an additional archived and Bzipped file will be present in the folder.

See below for list of folder contents

<i><b>Folder</b></i>	<i><b>Folder contents</b></i>
<b>Output/</b> e.g 20140408_RCAS_17_18.4e_wt_rec/	Metadata/ Cropped/ Scaled stacks/ Compressed cropped folder: Cropped_[date]_[gene]_[stage]_[litter]_[zygosity]_[sex]_rec..tar.bzip
<b>Cropped/</b>	Cropped recon files Additional files from original recon
<b>Scaled stacks/</b>	A “tiff” stack for each scaling type. The name will include the scaling factor and the new pixel size. e.g: 20140408_RCAS_17_18.4e_wt_rec_scaled_x2_pixel_8.9626
<b>Metadata/</b>	config4user.txt configobject max_intensity_z.tif pid.log X.X_scale.log e.g 2.0_scale.log session.log

See below for further details about the metadata:

The **session.log** file is a general record of what processing was performed.

The **X.X\_scale.log** file is a record of what occurred in the scaling performed by ImageJ.

The **configobject.txt** is a Python Pickle object used to save the parameters HARP uses. The **config4user.log** is more user friendly readable version of the Pickle file.

The **cropbox.txt** file is Python pickle file of the cropping dimensions used.

The **max\_intensity\_z.tif** is the image used for manual cropping using the “Get dimensions” option.

## Appendix

### 7-Zip instructions

7-Zip is a free and open source compression and archiving tool for Windows.

**Download from:** <http://www.7-zip.org/download.html>

#### To un-archive and unzip:

7-Zip should be accessible from right clicking on the document. The user can navigate to “Extract here”. This will unzip the file.

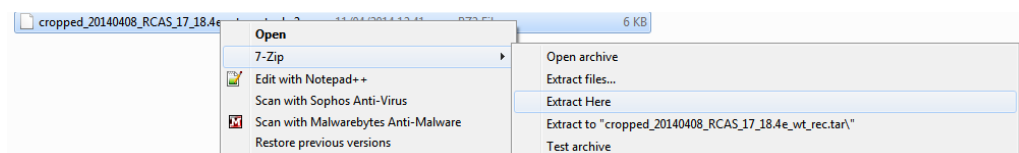


FIGURE 12: 7-Zip first step.

The user should then right click and “Extract to ...” to un-archive the folder.

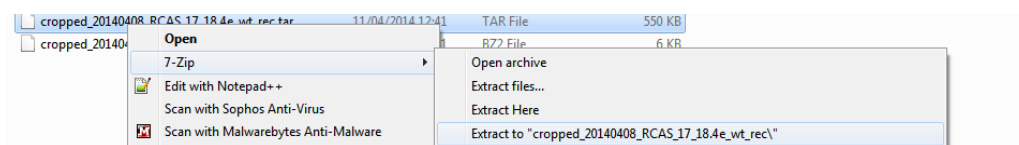


FIGURE 13: 7-Zip second step.

Alternatively 7-Zip file manager can be opened up and the steps shown previously can be carried out inside the browser.

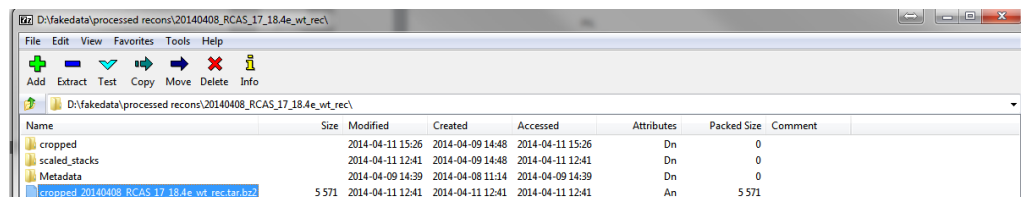


FIGURE 14: 7-Zip file manager



## **Required external python modules**

Matplotlib

Numpy

psutil

PyQt4

... todo