

# Image Analysis with IAP

Example Workflow for Analysis of Barley Images - exported using the 'DBImportExport'-Tool.

Dr. Christian Klukas, Head of Image Analysis Group, Leibniz Institute of Plant Genetics and Crop Plant Research (IPK) Corrensstr. 3 Gatersleben, Work State 06466 T 039482 5 763 F 039482 5 407 Work Email klukas@ipk-gatersleben.de

Document revision 2: November 2013, updated to reflect changes, introduced with IAP V1.1.

# Data set export and loading

### Export image, watering and weight data from the LemnaTec database

At first, the data set needs to be exported from the LemnaTec system. This can be done using two approaches. The first one requires the setup of an direct database connection of the IAP system to the LemnaTec database and the according FTP server. As this requires additional detail knowledge and according access information, this procedure is described in the IAP system documentation (available from within IAP using the command sequence **Start > About > User Documentation**.



The second approach, described here in more detail, is the export of the data using the LemnaTec

**Lemna DBImportExport** - Tool. The startup icon of this program is situated on the Windows Desktop, once the program has been installed.

- 1. Start the Lemna DBImportExport program.
- 2. Click Database-Login.
- Select your user name (if not available, ask another user with Administrator rights to create one for you) and enter your password. Select the correct database, which contains the desired data (in this example *BGH\_12*). Then click OK.

At IPK database names start with either "APH", "BGH", or "CGH", which means "Arabidopsis Phytochamber", "Barley Greenhouse" and "Corn Greenhouse", respectively. After that the relevant year (last 2 digits) is mentioned. Then click 'OK'.



📓 Login	
<b>S</b>	emnaTec
Benutzername [ Passwort: [ Datenbank: [	klukas 🔽 ••• BGH_11 🔽
	OK & Abbrechen

- 4. Click Snapshot-Export.
- 5. Click Dump export.
- 6. Click Snapshot selection.
- 7. Select the desired measurement label (name of experiment). In this example, select **1248KN**.

Then click the big **right-facing arrow** button to add the relevant snapshot list.

If you would like to work with an additional experiment, it is possible to select another measurement label and to add this to the list as well at this point.

#### Then click **OK**.

At the IPK measurement labels need to be named according to this systematic: Year Week Short Name \_ optional Description. For example an experiment of Kerstin Neumann, within the year 2012, week 48 would be named "1248KN". If additional information needs to mentioned, it could be named "1248KN\_drought\_stress". Special care should be taken, to not include special characters in the measurement names or in the plant IDs (during experiment setup or meta data import).







🕼 Snapshots auswählen				
11286 Snapshots wurden zu Ihrer Auswahl hinzugefügt.				
Snapshots in der Datenbank		Ausgewählte Sn	apshots	
🔗 barra a 🕅 🖂 🚱 🧶 waa 🛛 🕅	N	ID-Tag V	Zeitstempel	Messreihenbezeichnung 🔺
Sortiert nach: 📖 🧭 🖏 Weitere 🔹 📋 🔹		1036CS017	2011.05.05	1107KN
Navigation: 🕊 🗶 🔉 🛸		1051CS039	2011.05.05	1107KN
		1107KN001	2011.02.19	1107KN
Sortiert nach: Messung (85163)	4	1107KN001	2011.02.20	1107KN
Messung: 110/KN (11286)		1107KN001	2011.02.21	1107KN
Messung: 1116KN (6/81)		1107KN001	2011.02.21	1107KN
Messung: 1121KN (17993)     Messung: 1121KN (17993)		1107KN001	2011.02.21	1107KN
Hessung: 1122KN (142)     Hessung: 1122KN (142)		1107KN001	2011.02.22	1107KN
		1107KN001	2011.02.23	1107KN
🖶 – 🥔 Messung: 1137KN (18072)		1107KN001	2011.02.24	1107KN
Messung: 1142KN (12544)		1107KN001	2011.02.25	1107KN
🗄 🍃 Messung: Foto (5)		1107KN001	2011.02.26	1107KN
Messung: Topfblau (2)		1107KN001	2011.02.28	1107KN
Hessung: unknown (9)		1107KN001	2011.02.28	1107KN
		1107KN001	2011.03.01	1107KN
		1107KN001	2011.03.02	1107KN
		1107KN001	2011.03.03	1107KN
		1107KN001	2011.03.04	1107KN
		1107KN001	2011.03.05	1107KN
		1107KN001	2011.03.06	1107KN
		1107KN001	2011.03.07	1107KN
		1107KN001	2011.03.08	1107KN
		1107KN001	2011.03.09	1107KN
		1107KN001	2011.03.10	1107KN
		1107KN001	2011.03.11	1107KN
		1107KN001	2011.03.12	1107KN
		1107KN001	2011.03.13	1107KN
		1107KN001	2011.03.14	1107KN
		1107KN001	2011.03.15	1107KN
		1107KN001	2011.03.17	1107KN
		1107KN001	2011.03.18	1107KN
		1107/0001	2014 02 24	1107VN
			III	•
				OK Abbrechen

- 8. Click **Select output folder** and choose your target folder in the file system.
- 9. Click **No** in the **radio button** selection 'Combine images'.

Now the 'Start' buttons becomes enabled.

#### Click Start.

10. The relevant snapshots (image data, weight measurement and watering data) will be exported.

In case of intermediate errors, it is possible to restart the export, from within this progress dialog, once it is finished. It is also possible to close the program and later repeat the procedure. Existing data will not be exported again.

At the IPK the data sets from previous years may be archived to the so called Hierarchical Storage Management (HSM). In this case the export may take an extraordinary long time or even produce errors. In this case contact AG BIT or send an "DV-

Auftrag", to let the relevant data be "staged" on hard disc. Once this process is completed, the data may be accessed with normal speed.



SIM DBImportExport@BGH_11 (de-DE)		×
Snapshot-Dump-Export	G	urück
Snapshots auswählen	11286 snapshots selected	
Ausgabe-Ordner wählen	E:\LT-Export\1107KN	
Gekachelte Bilder kombinieren	C ja 💿 nein	
Start		
	🕜 Zurücksetzen	

DBImpo	rtExport@BGH_11 (de-DE)	x
2/11288: Cr 2/11288: Ali 3/11288: Du	<pre>scking directory 'E:\L1-Export\110/KN' for present snapsho sady dumped snapshots found: 0 mping 1/11286</pre>	ts.
4/11288: Du	mping 2/11286 mping 3/11286	
<b>I I I I I I I I I I</b>	III	•
STOP	Beenden	
5	Zurück	

# Preparation of meta data file

For proper operation of the LemnaTec imaging hardware, it is required to import a meta data file, which contains a list of plant IDs and their mapping to a particular measurement label. Also the so called "overall configuration" (OAC) needs to be specified in this list.

Additional information about the particular species, variety, treatments or growth conditions can be added in this table, too. Within the LemnaTec control software this extended annotation information is not processed, but it may be useful at later stages, e.g. when accessing the data with IAP.

The work with meta data is greatly simplified, if common identifiers and names are used. It is strongly recommended to the naming standards, described in the file "LemnaTec Operation Standards" by C. Klukas and I. Mücke (<u>http://ba-13.ipk-gatersleben.de/standards.pdf</u>).

It is recommended to add the following columns to the initial set of columns, required by the LemnaTec system (Barcode, OAC, Experiment): "Species", "Genotype", "Variety", "Growth Conditions", "Treatment", "Sequence" and "SeedDate".

If the "SeedDate" columns is available and mapped within IAP to the annotation property "sequence", time values will not start with the first day on the system ("day 1"), but will be time-shifted according to the seed date, and will for example been shown as "DAS 7", for days after sawing 7. This makes comparisons easier.

#### An example meta data file is shown, here:

Barcode	OAC	Experiment	genotype	SeedDate	Variety	Species
1248KN001	B_Midplant_9	1248KN	4	2102012	wild	Barley
1248KN002	B_Midplant_9	1248KN	4	2102012	wild	Barley
1248KN003	B_Midplant_9	1248KN	4	2102012	wild	Barley
1248KN004	B_Midplant_9	1248KN	4	2102012	wild	Barley
			· ·			

This file needs to be imported using the LemnaTec control software ("Lemna Launcher") as an CSV file.

For import into IAP this meta data file can be formatted as CSV or as XLSX file. The meta data file needs to be placed in the folder, which was used during snapshot data export (see previous section):

1248KN_1248KN520_2013-01-17_11-48-05_260521	24.08.2013 20:48		Folder
1248KN_1248KN520_2013-01-18_11-11-37_266231	24.08.2013 20:48		Folder
1248KN_1248KN520_2013-01-19_01-41-53_272981	24.08.2013 20:48		Folder
1248KN_1248KN520_2013-01-21_02-00-04_276501	24.08.2013 20:48		Folder
1248KN_1248KN520_2013-01-21_07-56-21_274021	24.08.2013 20:48		Folder
1248KN_1248KN520_2013-01-22_10-18-01_283241	24.08.2013 20:48		Folder
1248KN_1248KN520_2013-01-23_11-12-10_288951	24.08.2013 20:48		Folder
1248KN_Meta.xlsx	04.03.2013 22:52	33 KB	Micror
M Drive B ▷ ishare2 ▷ ishare2 ▷ ishare2 ▷ ishare2 ▷ in LT export ▷ in 1248KN ▷ in 1248KN_Meta.xl	SX		

The name of the meta data file can be freely chosen, but the file extension and type needs to be either CSV or XLSX. If you export several different data sets (measurement labels) from the LemnaTec system, you should also place all of the relevant meta data table files in this folder.

# Loading of exported data and of meta data into IAP

Once the data is exported and the meta data file has been placed into the export-folder, the data set can be loaded with IAP, using the button command sequence **Start > Load or Create Dataset > Create IAP Dataset(s) from LT Dump** (IAP

V1.0: Start > Load Dataset > DB-Imp.-Exp. Dataset). Navigate to the target folder and click the **OK**-button. Depending on the size of the data set and the load of the system, this may take a few minutes.

Once the data within the export-folder has been processed, a list (or a single) experiment-button appears. Click onto the data set button (in this case **1248KN**), to investigate the data set in detail. Click **View/Export Data** and use the tree view on the left, to check the meta data assignment:

<ul> <li>♥ Wis (RCB_Side_Mid</li> <li>♥ ① 0: Barley/Ar</li> <li>♥ ② 0: Barley/E</li> <li>♥ ③ 0: Barley/F</li> </ul>	_Plant) ta/cultivated irke/cultivated arine/cultivated /wild /wild /wild /wild	
- C 0: Barley/F	Property	Value
<ul> <li>G: Barley/F</li> <li>G: Barley/F</li> </ul>	Species	Barley
🕶 📑 0: Barley/F	Genotype	
O: Barley/F	Variety	wild
► 📑 0: Barley/F	Sequence	
• 📑 0: Barley/F	Growth conditions	
← ☐ 0: Barley/F	Treatment	
• 1 0: Barlev/F/	wild	



If the meta data from the meta data table is not correctly assigned to the build-in meta data fields 'Species', 'Genotype', 'Variety', 'Sequence', 'Growth conditions', or 'Treatment', click **Start > Settings > Metadata > Columns**. For each column in the meta data file, a new command button is created upon first processing of the meta data file. These command buttons can be used to adjust the assignment to the build-in meta data fields:



If several columns from the meta data file are assigned to a single build-in meta data field, it may be useful, to add the column name as a prefix to the build-in storage field. The 'Yes/No' toggle buttons next to the radio-selection command buttons for each field can be used this way (e.g. **Region - Use As Prefix**).

After adjustment of the meta data mapping from annotation file to the build-in IAP meta data annotation fields, <u>the data set</u> <u>should be loaded a second time</u>, by going back and clicking the **Load Dataset** button in the command history bar at the top. Now click again **DB-Imp.-Exp. Dataset**, load the data set and confirm the proper assignment of the meta data fields. If you find errors, <u>repeat the described procedures until all relevant fields are correctly assigned</u>.

If you use the same meta data column headings in the future, this adaptation needs only be done for the first experiment.

# Copying the data set to a IAP storage location

The loaded export-data set needs to be copied an IAP storage location, so that changes to the data set can be stored. The data set can only be analyzed, once it is copied to a IAP storage location. IAP currently supports two kinds of storage locations: (1) a MongoDB data base storage location and (2) a file system (VFS, virtual file system) - based storage location. In this workflow example, only the second case will be considered, as the first type of storage requires the configuration of a MongoDB database. This is described in the IAP User Documentation (http://iap.ipk-gatersleben.de/documentation.pdf).

File based storage locations can be easily set-up as (in this example, we use the second option):

- A permanent stored storage point, by using the command sequence Start > Load or Create Dataset > Add Remote Storage Location > Local. By using this command sequence a specific folder button is created on the 'Start'-screen and available when selecting the experiment command button 'Copy'.
- Or on an ad-hoc basis, by using the Start > Load or Create Dataset > Load IAP Dataset(s) from File System command to open a storage folder and the Copy > To Local File System command button to place an data set in a folder. The latter command sequence is available from a loaded experiment data set.

In both cases a storage folder may be used to store multiple input data sets and also multiple analysis result data sets.

Within the example workflow, click the **1248KN**-command button (if not yet loaded). Now it is time to check the experiment header annotation: *Name, Coordinator, Experiment-Type*, and more:

● ● ●	IAP – The Integrated Analysis Platform	
Start	Load Dataset DB-ImpExp. Dataset Updated 7 months ago	
To Main Storage (BA	-24) To Local File System To Remote Storage Location	
Experiment 124	8KN	
Name	1248KN	
ID	lt:null:1248KN	
Import by	klukas	
Origin	lt:null:1248KN	
Database	null	
Coordinator	Neumann, Dr. Kerstin (GED)	
Access Group		
Experiment-Type	Barley 👻	
Start-Time	29.11.2012	
End-Time	23.01.2013	
Sequence/Stress		
Remark		
Outliers		
Connected Files	305.532 (0 KB)	
Numeric Values	333,465	
Storage Time	01.01.70 01:00	
History	-	
Analysis Settings	(not assigned)	
	Cancel Save Changes	

Modify the entry fields as desired and then click the **Save Changes** button. The changed settings will not be applied in memory and will be used during the copy operation. By changing these settings before the copy operation it is ensured, that the data is stored from the beginning in the right folder structure.

Then choose **Copy** and click **To Local File System**. Locate or create the desired storage location. This should be an initially empty folder! In this case you could create and select folder named *1248KN\_datasets*. Then click **Select Target Folder** within the file selection dialog.

Depending on the system storage capabilities and the data set size, the transfer may require up to a few hours or even days to complete (an more exact estimate is shown after completion of a larger part of the work).



# Loading the data set from a IAP storage location

Once the data set transfer is complete, it is time to load the data set from the storage folder: Click **Start > Load or Create Dataset > Load IAP Dataset(s) from File System** and choose the folder, which contains the IAP experiment data structures (the folders *index, conditions, data,* and *icons*), then click **Select Target Folder**.

Now click **Barley > Neumann, Dr. Kerstin (GED) > 1248KN** (the actual button text depends on the experiment header values for *Coordinator* and *Experiment Type*).

	IAP -	The Integrated Analysis Platform		
Start Lo	ad Dataset /Volumes/klukas/Barlomics Summer Schor	ol Tutorial/1248KN_datasets Barley (1)	Neuman, Dr. Kerstin (GED) (1)	248KN ed 8 sec ago
View/Export Data A	nalysis PDF Report Copy Tools	Add to Clipboard Put data in 'Experim	nents' tab Move to Trash	
✓ Experiment 1248K	N			
NameIDIDOriginDatabaseTordinatorAccess GroupExperiment-TypeStart-TimeEnd-TimeSequence/StressRemarkOutliersConnected FilesNumeric ValuesStorage TimeHistoryAnalysis Settings	1248KN  iser.dir.1377786773345:index/1377785487652_0_kl ilukas  iser.dir.1377786732730:index/1377785487652_0_kl iluli  Veuman, Dr. Kerstin (GED)  Barley  29.11.2012  23.01.2013  4,516 (4,616 MB)  4,810  19.08.13 16:32  . not assigned)	Jkas_1248KN.iap.index.csv Jkas_1248KN.iap.index.csv		

Vou may also at this point change the experiment header annotation; when finished click **Save Changes**. Some folder structures (e.g. the folders used for image storage) will not be updated, as a copy or move operation would require a longer time. The data set references in this case the images from the previous storage sub directory within the storage location.

# Image analysis

# Assignment of an analysis template

IAP V1 comes with four analysis templates: *Arabidopsis Analysis, Barley Analysis, Maize Analysis* and *Scanned Roots Analysis.* In our case the *Barley Analysis* analysis template is best suited to our data set. The first step is to select and assign this analysis template to our data set. Once the experiment data set is loaded, click **Analysis > Select Analysis Template**. Now click onto the command-button **Use Barley Analysis**:



You may now click the data set button 1248KN in the command history, to navigate back to that screen.

# Image browsing and selection of a test images

A pre-defined or user-customized analysis template should be tested with a selection of test images from the current data set, before the whole data set is analyzed. Click **View/Export Data**, to open the data browsing view, which consists of a tree view on the left hand side and a image display / test diagram plotting button selection on the right hand side:



On the left, the different camera sources and time points are placed into a tree view. At first you may select the side or top view. Then the camera types (fluo/nir/vis for fluorescence, near-infrared and visible light) can be opened. Then, the

experiment factors are listed (e.g. genotypes and treatment combinations). After that the experiment days are displayed.

Once you select either a experiment factor combination (called *Condition* within IAP) or a day, or by opening the day item, a image, on the right hand side the connected image set is displayed in a icon-view. As you move your mouse over these images, two checkboxes appear, one named **Outlier** and one named **Flagged**. *Outliers* are not analyzed during the image analysis. The attribute *Flagged* is currently not further processed within the application and thus can be freely used.



If all images from a certain day, until a certain day or from a certain day are outliers, or if plants with a specific plant ID, or if a plant from a certain day onwards are outliers, then this information can be specified in the experiment header in the field **Outliers**. Use the following encoding to specify such outliers:

>40	÷	Mark all images after day 40 as outlier
>=40	÷	Mark all images after and including day 40 as outlier
<10	÷	Mark all images before day 10 as outlier
<=10	÷	Mark all images before and including day 10 as outlier
PlantID	÷	Mark all images with plant ID X as outlier (e.g. 1248KN001)
PlantID/20	<b>~</b>	Mark all images with plant ID X from day 20 on as outlier

You may combine multiple specifications by using '// as divider. For example, you could enter "<10 // > 40 // 1248KN001 // 1248KN002', to mark all images before day 10 and after 40 as well as all images independently form time from 1248KN001 and 1248KN002 as outlier images.

To display a full size view of a specific images right-click the image and choose the context menu item **Show Image**. The image will be loaded and shown. In addition the ImageJ command window is shown (an open source image processing library, which is included within IAP). This command window contains useful manipulation commands, such as a magnifying tool (right-click the image to zoom-out, and left-click the image to zoom-in, once this command is selected). The ImageJ command window and additional output windows of ImageJ can safely be closed at any time.

The image display itself uses a color-code (replaced colors), to indicate pure white areas. These areas are shown much darker, which helps in filter calibration, as otherwise slightly gray bright spots in the image can hardly been seen and identified:



To test the *Barley Analysis* pipeline, right-click an image and choose **Barley Analysis (Image+Reference)**. After the analysis is completed, the following window is shown:



This image window (large pure black areas in the lower part and on the side have been removed for print-out), contains a list of images, representing the input and output of the individual analysis steps. On the lower part of the window, a scroll bar is shown, which can be used to navigate through the individual pipeline steps (the current step is shown at the top of the image).

If not all parts of this image window are visible, activate the ImageJ magnifying tool button command mode and right-click into the image view, to make the window view smaller.

#### **Pipeline settings adjustments**

As you scroll through the individual pipeline steps within the pipeline debug window, you should carefully check the results of the individual pipeline steps. Sometimes, depending on illumination settings, different plant treatments (e.g. severe drought stress), or other factors, a pre-defined pipeline yields to unsatisfactory results. In this case most importantly check the individual pipeline block-settings of the following analysis blocks:

1.	Align	÷	Adjust image position and then size, to prepare proper alignment for overlay $% \left( {{{\boldsymbol{x}}_{i}}} \right)$
			of camera images. If possible, make sure that the camera settings and positions are properly adjusted. If this is the case, the settings for this block need only very little fine tuning.
2.	Detect Blue Markers	÷	Settings to enable conversion of pixel values (size, height) to mm.
3.	Lab-Filter	÷	Color filter, to remove white/gray background, blue support frames, soil
4.	Intensity Calculation F.	÷	Fluorescence intensity calculation and filtering.

Þ

5.	Remove Small Objects	÷	Remove small remaining disturbing elements from visible light and
			fluorescence images.
6.	Use Fluorescence Mask	÷	Overlay masks to remove elements within visible light and fluorescence
	to Clear Others		images, which are not part of the other camera type. Use fluorescence mask to clear parts of the infrared image. The <b>Align</b> block settings should be properly selected, to avoid necessity for too large mask-blur-factors.
7.	Cut From Side	÷	Remove large parts from the top, bottom and side, e.g. side doors in the
			images.
х.	Remaining Blocks	÷	The settings of further blocks may be adjusted, detailed information
			for each block is part of the user documentation.

To adjust the individual pipeline block settings, click the button **Change analysis settings** in the pipeline debug window. This button is shown below the image display. If it is not visible (in case of smaller screen resolution), select the ImageJ zoom toolbar icon and *right-click* into the image display, the window will become smaller and the full window including the two command buttons at the end of the window becomes visible. Depending, on whether you are currently optimizing the pipeline for the analysis of *Side* or *Top* images, the following window appears:

O     O     Change analysis settings (experiment 1248KN)	
Modify settings Side Settings	
Loadmanes Alian ColorBalanceVerticalFluo CreateDummyReferencelfNeeded ColorBalanceVerticalVis	Defaults
	(delayed)
ColorBalanceVerticalNir ColorBalanceCircularVisNir RemoveBackground DetectBlueMarkers MedianFilter LabFilter	
Closing ikdiff intensityCalculationFluo ClearNirPotFromNir RemoveSmallObjectsVISFluo UseFluoMaskToClearOther	
Image: Constraint of the second se	
CutFromSide AdaptiveThresholdNir SkeletonizeNir CopyImagesApplyMask SkeletonizeVisFluo 1350.0	
LeafCurlingAnalysis CalcMainAxis CalcWidthAndHeight CalcColorHistograms CalcConvexHull RunPostProcessors	
DrawSkeleton MoveMasksTolmageSet Crop HighlightNullResults	

The shown buttons are ordered accordingly their execution sequence. In case pipeline blocks are removed or disabled, the according settings will *not* be removed. Settings for newly added blocks will appear at the end of the button list. If not all command buttons are shown (redraw/layout bug), slightly resize the window. If your screen resolution is not high enough to display all buttons, *right-click* into the empty button area at the top of the window and choose the context menu command **Compact (2)**.

#### 1. Adjusting the alignment of camera images (Block Align)

To adjust the settings of this block, select **Modify settings > Side[/Top] Settings > Align**:



Enable debug:



To get a visual feedback while adjusting the settings, click **Debug**, and then go back to the pipeline debug window and click the command button **Re-run Analysis (debug)**. While the analysis is repeated, close this pipeline debug window. Once finished, the results of the new analysis will be opened and displayed in a new pipeline debug window. The *Change analysis settings* window does not need to be closed and can stay open.

The *debug* function of the align analysis block enables the display of two overlay images with the title *VIS* & *FLUO* and *NIR* & *FLUO*. This screenshot shows the VIS & FLUO image:



Now place this window next to the *Change analysis settings* window. It is recommended not to change the fluorescence settings, but instead to only change the Vis - settings and the Nir - settings. In this case mainly a horizontal offset is apparent. The visible light image is shown slightly shifted vertically, below the fluorescence image. Use this scheme, to adjust the individual settings (first change the position, *then* the size):

VIS lower than FLUO	<i></i>	Decrease Vis Shift Y
VIS above FLUO	<i>(</i>	Increase Vis Shift Y
VIS left of FLUO		Increase Vis Shift X
VIS right of FLUO		Decrease Vis Shift X
VIS correctly centered but larger than FLUO		Decrease Vis Zoom X and VIS Zoom Y
VIS correctly centered but smaller than FLUO		Increase Vis Zoom X and VIS Zoom Y
VIS correctly centered but not as wide than FLUO		Increase Vis Zoom X
VIS correctly centered but wider than FLUO	<i>—</i>	Decrease Vis Zoom X
VIS correctly centered but not taller than FLUO		Increase Vis Zoom Y
VIS correctly centered but not as tall as FLUO	<i>~</i>	Decrease Vis Zoom Y

As you change a setting and click **OK**, the image display is immediately updated. Adjust the NIR position in the same way as the VIS position adjustment.

After adjusting the Vis Shift X value to 20, and the Vis Shift Y to -1, the image looks like this:

As soon as the results are satisfiable, disable the debug function of this block by clicking **Debug**. Re-run the analysis by clicking **Re-run Analysis** (debug) and proceed with adjusting the remaining important block settings.



#### 2. Adjusting the conversion of pixel-values to mm-values (Block Detect Blue Markers)

Two general cases need to be considered: (1) the visible light image contains blue marker spots at the left and right side, or (2) there are no blue marker spots in the image:

Case 1: Blue marker spots are available:

These spots will be detected and removed from the image, automatically (if **Remove Blue Markers From Image** is enabled and **Use Fixed Marker Distance** is disabled). The only setting to be adjusted is the real horizontal distance of these marker points. Click **Modify settings > Side[/Top] Settings > Real Blue Marker Distance** and enter the vertical distance in mm. This approach has the advantage, that zoom changes and even slight daily zoom variations are detected and used to calculate more accurate height and size information values (for mm-based measurements).

Case 2: Blue marker spots are not available:

In this case a fixed conversion can be used and specified. Click **Modify settings > Side[/Top] Settings > DetectBlueMarkers** in the *Change analysis settings* window. Now *disable* the setting **Remove Blue Markers From** 

**Image** and *enable* **Use Fixed Marker Distance**. Go back to the main IAP window (titled *IAP - The Integrated Analysis Platform*), and *right-click* onto a image button, choose the context menu item **Show Image**:

Then use the ImageJ rectangle-selection tool and click and drag the mouse over an object in the image display (titled *Main Image*).





(an object of known size should be marked)

The horizontal extend of the measured object (e.g. horizontal carrier dimension or door-to-door measurement) is shown in the ImageJ status bar next to the text 'w='. Enter the displayed value in the settings field **Fixed Marker Distance (px) (...)**.

Now click **Modify settings > Side[/Top] Settings > Real Blue Marker Distance** and enter the actual real-world size of the just measured object in mm. So in this case *two* values need to be entered: the size of an object in pixels (from the taken image) and the size of the object in real life in mm. From that the measurements in px, can be converted to mm already within the IAP program.

You may also opt to skip this step and just disable this block and convert the needed values during the post-processing by yourself.

#### 3. Adjusting the color-based foreground/background separation (Block K-Means-Filter)

With IAP V1.1 the K-Means-Segmentation-block has been introduced and is used by default by most of the included example analysis templates.



To access the settings of this block click **Modify settings > Side[/Top] Settings > Auto-tuning VIS-Segmentation (K-Means)**. The following parameters can be adjusted:

Color Classes	÷	Number of defined color-classes
Seed Color X	÷	Color of the color class 'X' (X=[1n], n=Color Classes). By clicking the color-icon you
		may modify the according color-class color (e.g. specify a green color for foreground plant color, or blue, for a support frame, which should be assigned to the background.
Foreground X	<i>—</i>	For each color class besides the color of the class, it needs to be specified, whether
		the specific color belongs to foreground or to the background. All image pixels, which are deemed most similar to a background color class will be removed from the image.

To increase the number of color classes, modify the value 'Color Classes', reanalyze a test-image and then modify the newly appearing 'Seed Color' and 'Foreground' settings of the new color classes. To reduce the number of classes, decrease the 'Color Classes' value, only the first color classes, up to specified number of classes, will be used for segmentation. Settings for then unused color classes will remain in the settings list.

#### 4. Fluorescence intensity calculation and filtering (Block Auto-tuning Fluo Segmentation)

The purpose of this block is to perform two operations: (1) deriving intensity values from the fluorescence image, and (2) to establish a cut-off filter, to remove pixels from the image, which fall below a certain threshold, e.g. non-living objects such as pots, side door elements or the background noise. IAP V1.1 automatically derives a dynamic cut-off, based on intensity of the background-noise.

 For intensity calculation three intensity values (ICh, IPh, ICl) are calculated from the input image. These values will be summarized for the plant image and stored within the result data set as '*intensity.chlorophyl.average*', '*intensity.phenol.average*', and '*intensity.classic.average*' by the analysis block *BICalcColorHist* in the *Feature Extraction* phase of the image analysis pipeline. It should be stressed, that the relationship and validity in comparison to actual chlorophyl content, phenol content or photosynthetic activity is not yet confirmed or tested. Reference measurements have been carried out but have not yet been analyzed in detail. It is possible, that the calculation formulas will be tweaked or modified in the future. Currently these values can be seen as



informative relative values, to compare different genotypes or stresses, and if indicative, to give hints for further analysis using specialized equipment.

ICh: The first value is related to the reflectance color, related to the *red* reflectance. The more green and the more bright the plant appears in the visible light image, the higher this value ICH (see below) will be. The calculated trait is called *'intensity.chlorophyl'*.

The calculation formula is: ICh = (1 - distToRed(hue)) \* red\_intensity Depending on the pixel color shade, the *distToRed* function returns 0 for yellow pixels, 0.5 for orange and 1 for red.

IPh : This value is related to the yellow reflectance, apparent mostly for severely stressed leafs, not green plant organs (e.g. flowers). The more bright yellow the fluorescence reflectance appears, the higher this value will be. The calculated trait is called *'intensity.chlorophyl'*.

The calculation formula is:

IPh = 1 - (1 - distToRed(hue)) \* red\_intensity
IPh < 0.333 => IPh = 0
Depending on the pixel color shade, the distToRed function returns 0 for yellow pixels, 0.5 for orange and 1 for red.

ICI : This value is related to the red/yellow reflectance and the brightness.

The calculation formula is:

ICl = red\_intensity / (1 + green\_intensity)

The output value is furthermore scaled factor of 1.2121 (/ 0.825 in the code), it is contained in the formula to scale the output range and improve the value range resolution, as the individual pixel intensity is stored as an pixel gray level, ranging from 0 to 255, and stored in the Blue channel of the RGB result image of this analysis block.

Within the code these three formulas are used to calculate *three* gray scale images. The result images are then combined into a RGB image, where channel R contains the values of ICh, G the values of IPh and B the values of ICl.

2. The calculated values are subjected to a intensity cut-off-filter. With IAP V1.1 the newly introduced block 'Auto-tuning Fluo Segmentation', automatically establishes cut-off-filter values, based on automated cut-off procedure (details on the algorithm will be added to the IAP System Documentation for IAP V1.1). If the automatically and dynamically derived cut-off values prove to be inadequate in your testing, you may disable the block setting **Auto-tune**. After re-analyzing the test-image and opening the block settings again, the cut-off values may be modified according to the following description:

Pixels with intensities below a certain threshold a removed from the output. As the output image uses low values for

*high* intensities, and *high* values for *low* intensities, and furthermore the RGB image uses a scale of 0 to 255, the cut of value lays in the range of 0 to 255 as well.

To access the settings of this block click **Modify settings > Side[/Top] Settings > LabFilter**. The shown parameters are called **Minimum-Intensity-Classic**, **Minimum-Intensity-Chloro**, and **Minimum-Intensity-PhenoI**. These should be increased, if parts of the plant are removed. These values should be lowered, if non-plant elements such as pots, transport systems or other elements remain in the image. If the reflectance is more in the yellow range, the 'PhenoI' parameter can be lowered, first. **In this example all values are changed to 250.** 

#### 5. Removing small remaining objects (block Remove Small Objects Vis Fluo)

After color filtering and further processing some small disturbing objects may still be contained in the output images. The purpose of this block is, to detect all separated (not connected) objects in the image and compare its size (filled area and/or vertical/horizontal size) to cut-off values. Objects that are smaller than the thresholds will be removed.

To access the settings of this block click **Modify settings > Side[/Top] Settings > RemoveSmallObjectsVisFluo**. The following parameters can be adjusted (enable **Debug** to check the result in non-scaled output windows):

Dilation Vis	÷	If there are small 'cuts' within the image, e.g. by
		thin plant support elements, in a pre-processing step a dilate operation may fill gaps of the given size, so that these plant elements become connected before the size-filter is active. The result plant pixels will be modified by this filter, but that change will later be corrected by the block 'BICopyImagesApplyMask'.
Noise-Size-Vis-Area	÷	Objects smaller than the given size (filled area of the object)
		will be removed.
Noise-Size-Vis-Dimension-Absolute	÷	If the 'Use Vis Area Parameter' is enabled, objects with a
		vertical or horizontal extend below the given threshold will be removed. The purpose of this optional filtering step is, to remove thin horizontal or vertical objects.
Use Vis Area Parameter	÷	If enabled, the vertical and horizontal object extend is
		evaluated (see previous parameter).
(corresponding Fluo settings)	÷	Also the fluorescence image may be subjected to size-
		based filtering. Normally, these values should be smaller than the visible light images, as less noise is apparent

light images.

An example input and output may look like this:

In this case, small disturbing elements from the pot-area are removed, but also larger parts of the plant itself.



in these images and sometimes also the camera resolution of the fluorescence images may be lower than the visible By using the shown parameter values, the output image retains most of the plant and removes a larger part of the noise from the plant pot.

Hint: Not all noise elements need to disappear, the fluorescence image will be applied to the visible light image at a later stage within the pipeline.

After adjusting the parameters *disable* the **Debug** function of this block.





#### 6. Use Fluorescence Mask to Clear Other Images (Block Use Fluo Mask to Clear Other)

A individual filtering step normally can't remove all of the noise elements of an image set, therefore different not 100% effective filter steps need to be combined, to achieve an improved overall filter quality (removal of most noise objects while removing as little as possible from the plant). As the fluorescence image normally contains the smallest number of disturbances, it is often a good idea to remove elements within the visible light image and the near-infrared image based on a mask, derived from the fluorescence image. Problematic are plant movements while taking the individual images. Therefore, normally no 100% match can be established for Barley or Maize plant images. But for *Arabidopsis*, often a near-100% match can be established, which simplifies the foreground/background separation to a large extend.

In all cases, the *alignment* of the images of the different camera systems should be properly adjusted (see step 1 of this optimization workflow description).

To access the settings of this block click Modify settings > Side[/Top] Settings > UseFluoMaskToClearOther.



In order to not remove slightly moved plant elements, the *blur*-factor values should be adjusted. Increase the **Blur Fluo Mask On Vis** value until the overlay of the fluorescence image causes no removal of plant parts in the visible light image. Vise-versa, increase the **Blur Fluo Mask On Vis** value, so that no part of the fluorescence image is removed as the resulting visible light image mask is applied back to the fluorescence image.

After adjusting the parameters disable the **Debug** function of this block.

#### 7. Remove larger parts at the image border areas (Block Cut From Side)

Sometimes larger parts of an image may easily be removed, e.g. parts below the plant pot, or larger image chamber side elements, e.g. doors or other construction parts.

#### To access the settings of this block click Modify settings > Side[/Top] Settings > CutFromSide:



This block provides two different operation modes: (1) Blue marker based cutting and (2) absolute cropping.

(1) For the marker-based cropping (default) an offset from left and right (**Marker Cut Left And Right Offset (percent)**) can be specified. The larger this value is set, the more space is left from left and right for the cutting. Negative values cut the image right of the left marker and left of the right marker, making the image smaller than the marker positions. The offsets are specified as relative values in percent of the width or height of the image. The same relative cut-off positions are used for all camera types, even so only markers within the visible light image are recognized.

(2) Disable **Crop (marker-Based)** to activate mode 2 which is cropping selected images based on the image borders. For demonstration purposes the **Cut Vis** setting has been enabled. Then perform an re-analysis (pipeline debug window: button **Re-run Analysis (debug)**) to create the non-marker based settings. Then click **Side Settings > CutFromSide**, to update the list of setting buttons. The settings screen contains the new settings **Cut-Off Vis From Left (percent)**, **Cut-Off Vis From Right (percent)**, ....from top, ....from bottom:

#### **Concluding words**

With this step the pipeline optimization is finished. The overall filtering and foreground / background separation is the result of several filters. Optimization may require adjustments of several blocks and is complicated by the fact, that illumination, plant colors and other factors can change over the course of the experiment time. Therefore it is advised to use several testimages (and to optimize the settings also



for the top-view images). Eventually, if there are major configuration shifts, it is recommended to tune the analysis separately for different time spans. Click 'Modify settings > Separate Settings', to split the settings into different time spans, e.g. using a separate set of settings for early or late time points, or if different camera configurations are used within the experiment, to use different pipeline settings for each camera configuration

# Starting the analysis (test run)

As you find that the analysis of selected top and side view images gives satisfactory results, you may proceed further to the goal of an overall analysis of the whole data set. It is recommended to cover the analysis of a wide range of the experiment time, as sometimes plants show different colors at the end of an experiment or at the beginning. Or it is sometimes useful to check as many plants as possible for a selected day.

First, make sure, that the **Debug** options of all blocks are disabled, and that all blocks are **Enabled**, which should be part of your analysis pipeline. Check the side and top settings by looking at the icons:

In this case, the **Align** block has the 'Debug'option enabled, and the **MedianFilter** block is not enabled. These or similar problems should be corrected.

Now close the *Change analysis settings* window and all other debug windows. Open another IAP interface window, by *right-clicking* into the empty area next to the command history buttons at the top of the IAP interface, then choose **New Window**:





In the new IAP window (window title 'IAP Cloud Storage, Analysis and Visualization System'), click **Start > Settings > Pipeline-debugging** and then enable the option **Debug-Only-Single-Plant**. In the first IAP interface window, with the image button browser visible, choose and note an plant ID name (e.g. *1248KN512*). Go back to the second IAP interface window, showing the pipeline-debug settings, and click **Debug-Single-Plant-Id**. Enter the desired plant ID. Go back to the first IAP interface window, in the command button history at the top click onto the experiment button, in this case **1248KN**. Then click **Analysis** and then **Perform Barley Analysis**.

If the ImageJ window is still open, you can close it (confirm the exit warning). The second IAP interface window can stay open or may be safely closed. Before closing an IAP interface window always check the title or window icon. The main IAP window has the title 'IAP - The Integrated Analysis Platform' and (not immediately visible on Mac systems) an picture of the earth as its window icon. Closing the main window causes IAP to quit without further notice.

During analysis the analysis button shows the progress. In this case only one plant ('1/1, plant 5;1248KN512') is analyzed. The analysis time of a single snapshot (consisting of a set of camera images) is also shown in the status area, e.g. '69s, 58s, ...'. The last line shows the amount of time, that is left to complete the analysis and the amount of time the analysis will overall take. Finally, the date and time of the estimated finish time is shown.

0	O     O     IAP - The Integrated Analysis Platform						
ļ	Start L	Local File Syster /Volumes/klukas/l	n Barlomics Summer School Tutorial/1248KN_datasets	Barley (1)	Neuman, Dr. Kerstin (GED) (1)	► L248KN updated 5 days ago	Analysis
5	Select Analysis Templ	late Export/Modify settings	Perform Barley Analysis = 		-		
	<ul> <li>Experiment 1248</li> <li>Name</li> <li>ID</li> <li>Import by</li> <li>Origin</li> <li>Database</li> <li>Coordinator</li> <li>Access Group</li> <li>Experiment-Type</li> <li>Start-Time</li> <li>End-Time</li> <li>Sequence/Stress</li> </ul>	KN [1248KN user.dir.1378193080735:index/13 klukas user.dir.1377786732730:index/13 null Neuman, Dr. Kerstin (GED) [ Barley [29.11.2012 [23.01.2013] [	377785487652_0_klukas_1248KN.iap.index.csv 377785487652_0_klukas_1248KN.iap.index.csv				

# After the analysis is completed, the result is saved to the file system location (or the MongoDB database, if the data set is loaded from there).

If the analysis is interrupted or does not complete correctly, the incomplete numeric data is not available any more. The file system or the database may contain intermediate result images. Consult the IAP User Documentation if you want to clean-up the image storage within the database. The file system storage location would need manual work for clean-up, which is not yet documented.

#### The analysis result data set is automatically loaded after completion of the analysis:



It can also be loaded newly, in this example by clicking Local File System > Analysis Results > [Coordinator Name] > Barley Analysis of 1249KN:



#### Click **View/Export Data**, to investigate the analysis results over time:



If problematic images have been found, write down the plant ID, time point and side angle of these images and go back to the original data set and proceed with the optimization as described in the previous chapter. If you just would like to follow or understand the analysis result, *right-click* onto an image button and choose **Barley Analysis (Reference + Old** 

#### Reference).

If the results are fine for the selected test plant, you may continue in analyzing the whole dataset. As described in the beginning of this section, open or go back to the second IAP interface window, click **Start > Settings > Pipeline-debugging** and *disable* the option **Debug-Only-Single-Plant**. Close this second IAP interface window.

### Starting the analysis (full run)

Now open the original dataset in the main and first IAP interface window, by clicking Local File System > Barley > [Coordinator Name] > 1248KN. Choose Analysis and click Perform Barley Analysis. Now the whole data set will be analyzed (in this example a sub-set of only 6 plants):



### Loading and export of the analysis results

After completion of the full analysis, the data set is saved and should be opened by clicking Local File System > Analysis Results > [Coordinator Name] > Barley Analysis of 1249KN.

Click **View/Export Data > Export Numeric Data** to access the export options for numeric data (spreadsheet files). To export the whole data set in form of an Excel file, click **Create Spreadsheet (XLSX) (complete...)**. Sometimes color histograms or colors of different sections (see IAP User Documentation for details) are not needed, and would only increase the number of data columns, making the result table needlessly very large. In this case, click **Create Spreadsheet (XLSX) (no histogram, no sections)**. Exporting in XLSX format has the big advantage, that no additional formatting work for different number formats is needed, as is the case for the CSV format. Contrary to the XLSX export, the CSV file will contain the result data separated for each side view (and top-view) angle. The XSLX export uses the average of the side view angles. In the CSV export the average data is also exported, the corresponding angle-column contains the value '-720', in this case.

# Post processing

# Automated report generation

IAP contains a build-in report functionality. Using this report function requires the installation of the statistical programming language *R*. R should be found in the application program search path within your operating system. Executing 'R' in a Windows/Mac/Linux terminal should load R and should not produce an error message.

In addition the report functionality requires the installation of LaTex.

If correctly installed and configured, you can access the report functionality by clicking onto the analysis result data set button (e.g. 'Barley Analysis of 1248KN'). Then click **PDF Report**. Now a number of options are displayed:

The line diagram plots in the output may show different lines, separated by the overall combination of conditions (enable button 'Group by Condition'), and by further meta data annotations, e.g. by *Variety* or *Treatment*. Normally up to two meta data fields may be used for grouping (e.g. separation of lines in the line diagrams).

If no grouping option is selected, the overall average of the traits is plotted, e.g. the average growth curve for all plants.

By default not all calculated traits are contained in the report, enable the option **Create Appendix**, to add additional plots in the report output, in the section 'Appendix'.

For this example, the following settings are used:

In the following, toggle buttons are shown, which make it possible, to filter-out selected groups, or to select only a specific list of groups (click **Toggle setting** and then choose the desired groups). For example it would be possible to filter-out specific genotypes or treatments, if these meta data groups have been selected in the previous step.

Then continue, by clicking the most left button in the action button bar (titled '... (click here to continue)'):



The last option groups allows the selection of so called 'Overview

Properties', plotted and included in the beginning of the report. In most cases, these can be left enabled. The first option button of this group (in this case 'Show Varieties in a single diagram'), switches between the possibility of plotting the





different data set groups in a single diagram (default), or to plot each group (genotype, treatment, ...) in a separate diagram next to each other in a grid.

#### Finally, click Create PDF (click here).

In the following, automatically the numeric data of the experiment is exported as an CSV file and report-creation scripts (batch / bash script files) and R code for plotting the data is saved on disc. Also the report

LaTex templates are exported. The R script plots all relevant data columns according to the previously provided user settings and then calls 'pdflatex' to create the final PDF report file. Once finished, the PDF file is opened in the default PDF viewer of your system.

As the R scripts require a certain set of R-packages, at first a PDF error page may be generated, with instructions on how to install / update the required R packages. If no output at all is shown after completion, the R installation, the LaTex installation or the IAP export may be incorrect of faulty. It then advised, to look for any visible error or log message in the terminal window, where IAP has been started and to contact the authors of the IAP system. Additional work and error checking is required to properly detect the multiple sources of problems on the different operating systems and user's configurations.

Also, currently, not all calculated traits are full sorted and ordered into their respective report sections, and more trait documentation will be included in the report with future upgrades of the IAP system. If the report function works, it provides an easy and quick method to get an overview about the wide range of calculated traits and their development over time, for the different genotypes or treatments within your experiment.

# **Further post processing**

To investigate the result data set further and individually, it is recommended to export the data as XLSX (as previously described in section 'Loading and export of the analysis results'). Once opened, the quickest way to group, filter and plot selected traits is the 'Pivot' function of Excel.



🔍 IAP , 🛄 ,