Lightroom Duplicate Image Finder

Scan your Lightroom image catalogue to

find suspected duplicates.

Introduction 1
Contact and Support 1
Requirements 1
Legal Stuff 1
What's Included 1
New Commands 1
New Smart Collections 1
Installation 2
Windows 2
Mac 2
Setting up Lightroom 3
Searching 4
Finding your Results 6
Known Problems 6
Smart Collection May Not Update (LR2) 6
Out of Memory Errors (LR2) 7

INTRODUCTION

This is a straightforward plug-in which will attempt to find duplicate images in your Lightroom catalogue.

CONTACT AND SUPPORT

You can get support or ask questions in the official helpdesk:

http://www.lightroom-plugins.com/contact.php

There are several FAQs available at:

http://www.lightroom-plugins.com/faq.php

REQUIREMENTS

Lightroom 2 or higher. Tested against Lightroom 2, 3 and 4.

LEGAL STUFF

This plug-in is not specifically endorsed by Adobe Inc®.

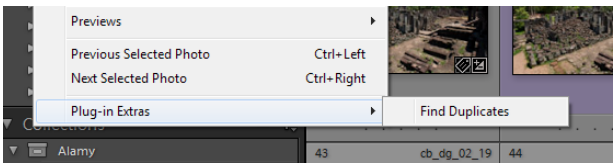
WHAT'S INCLUDED

NEW COMMANDS

One new command has been added to the "Library" menu in Lightroom:

1. ***Find Duplicates***

This performs the scan of the catalog, marking any suspected duplicates as it goes.



NEW SMART COLLECTIONS

To make it easier to find your images, there is a new Smart Collection provided. Smart Collections automatically find images in your catalog without having to be manually refreshed when something changes.

- ***Duplicate Pictures***

Shows all images which have been identified as being potentially duplicates.

INSTALLATION

This plug-in adds a new field to your Lightroom database. This is used to store a number indicating which images are seen as duplicates. Each set of duplicates will have a different number.

Before you install this, you should make sure you have a backup of your Lightroom catalog. If you do not already have automatic backups enabled, you can perform a single backup by going to "Edit→Catalog Settings…" and changing the "Back up catalog" list to "Next time Lightroom starts only", and restarting Lightroom.

WINDOWS

AUTOMATIC INSTALLATION

The easiest way to install on Windows is to use the MSI installer. Simply double-click and run through the prompts that appear.

MANUAL INSTALLATION

If you prefer to install manually, download the ZIP file instead of the MSI. When it has downloaded, unzip it and copy the folder called "`FindDupes.lrplugin`" into Lightroom's "`Modules`" folder. This can be found at:

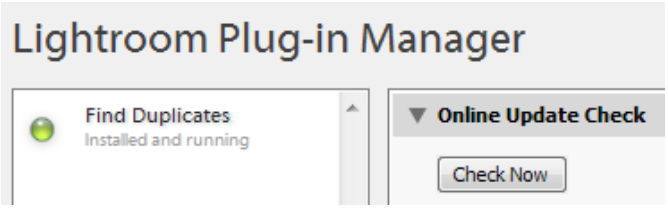
Windows XP: `%APPDATA%\Adobe\Lightroom\Modules\`

Vista/Windows 7: `%APPDATA%\Roaming\Adobe\Lightroom\Modules\`

MAC

Download the ZIP and extract it to your desktop. You should see a folder called "DuplicatesPlugin", containing the plug-in and a folder called "Smart Collections". You can place the plugin anywhere you like. Open Lightroom's Plugin Manager from the File menu, and click the "Add" button at the bottom of the list. Use the Finder window to select the plugin.

It should now be listed in the Plugin Manager as "Installed and Running".

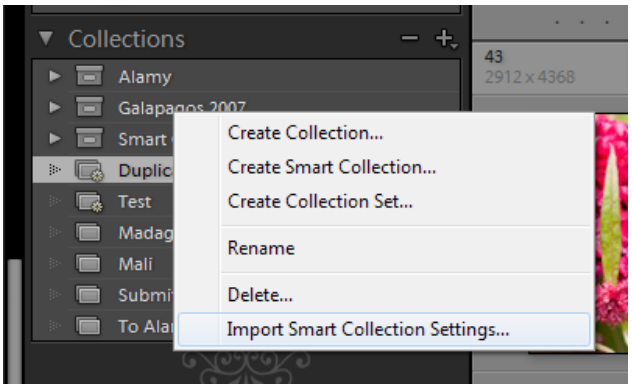


SETTING UP LIGHTROOM

INSTALL THE SMART COLLECTION (LIGHTROOM 2 ONLY)

You should install this for the plug-in to work. Since the duplicate images are marked using a custom field, they must be searched for and a Smart Collection is the simplest way of doing this. Installation is fairly straightforward:

1. Right-click inside the Collections panel and select "Import Smart Collection Settings".



2. Select the Smart Collection file. This is in a folder called "`Smart Collections`". If you installed using the ZIP file, it will be wherever you extracted the ZIP to. If you installed with the Windows MSI file, it will be in "`C:\Program Files\Jim Keir\Lightroom Duplicate Finder\Smart Collections`" or "`C:\Program Files (x86)\Jim Keir\Lightroom Duplicate Finder\Smart Collections`" if you have a 64-bit version of Windows.

If you are using Lightroom 3 or higher, the plugin will create a Smart Collection called "Found Duplicates" whenever you run it. There is no need to create the Smart Collection manually.

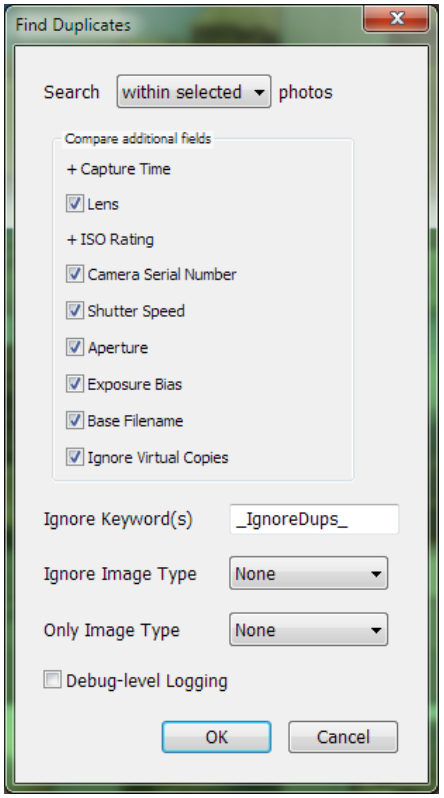
ENTER YOUR LICENCE KEY

If you have bought the plug-in, you should receive a licence key in an email. Without this, the plug-in will stop after finding a small number of duplicates. To enter the licence key, simply copy and paste it and your e-mail address from the email to the box in the Plugin Manager when the Find Duplicates plugin is selected.

SEARCHING

This assumes you already have some images stored in Lightroom.

Select the "Library→Plug-in Extras→Find Duplicates" menu item. A dialog will appear:



The plugin can search in one of three ways:

- ***All photos***

Your entire catalog is searched for duplicates.

- ***Within Selected***

Only duplicates within the currently selected photos are found.

- ***For Selected***

Duplicates of the selected photos are searched for in the entire catalog.

Next you can choose what EXIF data to compare, to determine if one photo is a duplicate of another. The top four are always used while the last four are optional, although they are likely to be always switched on. You could use these to decide whether to consider bracketed shots to be duplicates or not, for example.

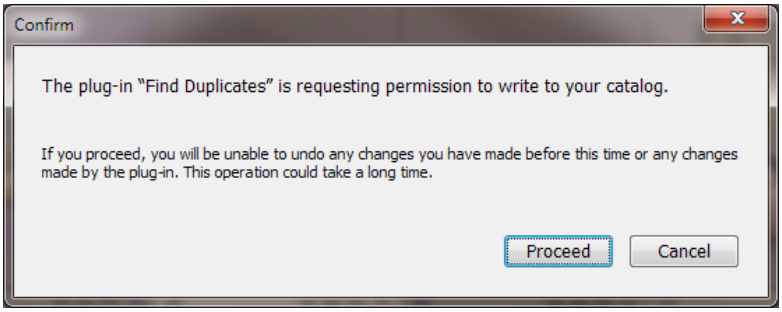
The exception is the final option, "Base Filename". All the other items are attributes of the photo itself and so are good candidates for deciding whether two photos are duplicates. The filename could easily change though, so you will need to decide whether to match filenames or not.

FILTERING YOUR RESULTS

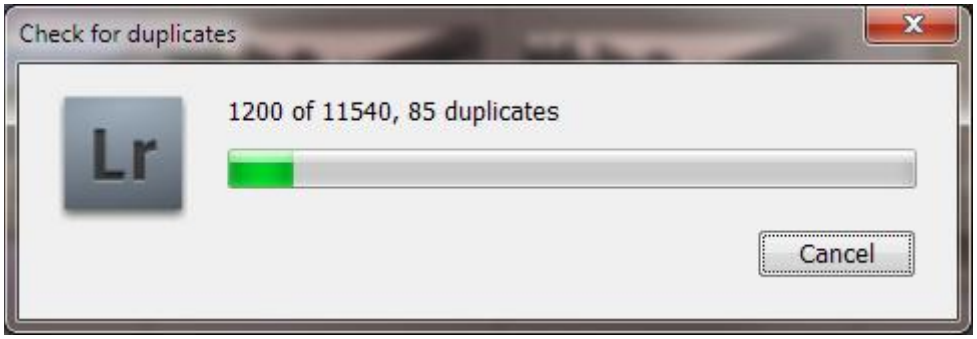
You can choose to ignore photos in several ways. One way is to have a specific keyword (or list of keywords) and to tell the plugin to ignore any photos with this keyword. For example, if you have uploaded processed copies of certain photos to an online picture agency, you can ignore all these processed copies by adding a non-exportable keyword to them and entering that keyword here.

Another way is to use the "Duplicate Search" option which can be found in the Library's Metadata panel under "All Plugin Metadata". By setting this to "Ignore", the duplicate finder will never consider this image to be a duplicate.

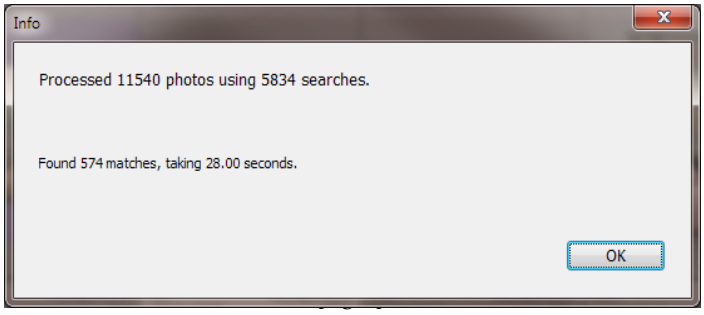
To begin the search, click "OK". A warning dialog will appear:



The only change to be written to the catalog is the unique identifier for each set of duplicates. Click "Proceed" to begin the search itself. A progress window will appear:



As a rough guide, a search of the 11,540 images in this catalog takes between 28 and 79 seconds on Windows running on an Intel i7-920, in Lightroom 2. Once the search has completed, a summary is shown:

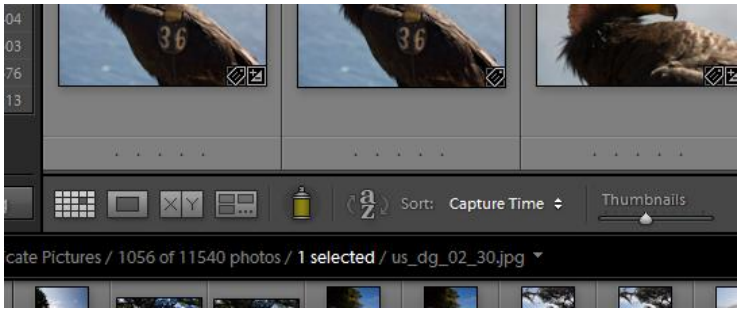


Take note of the number of matches found. Sometimes Lightroom ignores changes to values set by a plugin, and the Smart Collection will not update properly. Compare the number of matches in the dialog to the number reported in the “Duplicate Pictures” Smart Collection.

FINDING YOUR RESULTS

Once a search has completed, you can view the results. The search has filled in a new field in the database called “Duplicate Id”, and it is this that the Smart Collection searches on. Any photo which is considered to be a duplicate will have a number in this field, the same number being assigned to each duplicate of that photo.

To view the results, simply select the “Duplicate Pictures” Smart Collection. Unfortunately it's not possible to sort the results by the “Duplicate Id” field, but the next best thing is to sort by “Capture Time”:



Since the capture time will be identical only for duplicate images, this works just as well. It's now up to you to decide what to do with your duplicates. You can, of course, use Lightroom's normal Library Filter tools to limit the selection, for example by file extension, so that you can delete unneeded photos.

KNOWN PROBLEMS

SMART COLLECTION MAY NOT UPDATE (LIGHTROOM 2 ONLY)

Sometimes the Duplicate Images Smart Collection may not update correctly after running the Find Duplicates command. There doesn't appear to be a pattern to this, and it usually works correctly. It is a problem with all Smart Collections which refer to data fields added by plug-ins.

Restarting Lightroom restores this Smart Collection to show the correct contents.

OUT OF MEMORY ERRORS (LIGHTROOM 2 ONLY)

You may find that Lightroom gives errors about memory shortages if you have a large catalog. Lightroom 2 seems to use an unexpectedly large amount of memory when performing searches, and there is no way to fix this from within a plugin.

If this happens, the best workaround is to use “Search for selected photos” mode. Select a number of photos, maybe starting with 1000 or 2000, and search for duplicates for those. When the search is complete, quit Lightroom 2 and restart it, then repeat the search for the next few thousand.