

# Good practices

Aaron Ponti

Fall 2024

## Contents

<b>1 Guidelines for Best Practices in Image Processing</b>	<b>1</b>
1.1 Treating images as data	1
1.2 Saving the originals	1
1.3 Making simple adjustments	2
1.4 Cropping is usually OK	2
1.5 Comparing Images	2
1.6 Manipulating the Entire Image	2
1.7 Filters degrade data	3
1.8 Cloning degrades data	3
1.9 Making intensity measurements	3
1.10 Lossy Compression Degrades Data	3
1.11 Issues with Magnification	4
1.12 Issues with Pixels	4

This section discusses some of the best practices in scientific image analysis from the Online Learning Tool for Research Integrity and Image Processing (<https://ori.hhs.gov/education/products/RlandImages/default.html>) by the University of Alabama at Birmingham.

Please notice that there is much more detail on the website than we can treat here.

## 1 Guidelines for Best Practices in Image Processing

Each of this guidelines is accompanied by a video.

See [https://ori.hhs.gov/education/products/RlandImages/guidelines/photoshop\\_videos/default.html](https://ori.hhs.gov/education/products/RlandImages/guidelines/photoshop_videos/default.html).

### 1.1 Treating images as data

Scientific digital images are data that can be compromised by inappropriate manipulations.

- Each pixel in a digital image has:
  - a specific **location** relative to its neighbors,
  - a **scale** (size in microns) relative to the acquisition system used, and
  - an **intensity value** proportional to the amount of signal (light) that was captured.
- Because images are a **grid of numerical values**, they share many characteristics with a **spreadsheet**.
- Any **image manipulation** is simply a mathematical function that changes the underlying numbers in the image.
  - When we manipulate an image in Fiji or any other package, we are changing the underlying numerical values!
- Always **document** all image manipulations for **validation** and **reproducibility**!

### 1.2 Saving the originals

Manipulations of digital images should always be done on a copy of the raw image data. The original (including any associate metadata) must be retained.

- Processed images can be compared to unprocessed images to ensure that:
  - no important data was **lost** from the original images;
  - no **artifacts** were inadvertently added to the processed images.
- A copy of the unprocessed image is an extremely important protection against accusations of scientific misconduct.
  - Journals increasingly require that the authors provide raw data in the case of questions about the image processing or data interpretation during review.
  - Keeping a copy of the raw image data file is also a user's greatest protection against an inadvertent mistake in image processing.
  - The raw image data may initially be created in a vendor's proprietary file format. This formats usually contain metadata that describes the image acquisition conditions.
    - \* the original file format allows repeating the acquisition under the same conditions.
- When converting image data from a proprietary format, save your images as **TIFF** (Tagged Image File Format):
  - TIFF is a loss-less file format that supports a wide range of bit-depths.
  - Do **not** use the JPEG file format. JPEG uses a lossy form of compression and is inappropriate for most scientific images.

### 1.3 Making simple adjustments

- Adjustments like linear histogram stretches for better contrast are usually appropriate manipulations for digital images.
  - If the image is to be used for quantitative measurements, carefully document the changes.
- Avoid using software tools that make unclear, automatic adjustments to the images.

### 1.4 Cropping is usually OK

Cropping an image is usually acceptable.

- Cropping is usually used to remove empty background areas at the edges of the image.
- The act of selecting a specific area on a sample and the magnification at which to capture the image is a form of cropping. Be careful to avoid bias, such as selecting images that represent what you think your experiment sample should **look like**.
- Legitimate reasons for cropping include:
  - centering an area of interest;
  - trimming "empty" space around the edges of an image;
  - removing a piece of debris from the edge of the image.
- Questionable forms of cropping include removing information in a way that changes the context of the image; for example:
  - cropping out dead or dying cells, leaving only a healthy looking cell;
  - cropping out gel bands that might disagree with the hypothesis being proposed in the paper.

### 1.5 Comparing Images

Digital images that will be compared to one another should be acquired under identical conditions, and any post-acquisition image processing should also be identical.

- Several types of instruments, like confocal fluorescence microscopes, allow users to adjust the signal amplification settings on every image.
  - Use the same acquisition settings for all images to be compared.
- When a group of images is to be compared to one another, the processing of the individual images should be identical.
  - If there is a valid reason why the images in a figure were processed differently, the author must explain the rationale in either the methods section or the figure legend.

### 1.6 Manipulating the Entire Image

Manipulations that are specific to one area of an image and are not performed on other areas are questionable.

- Most forms of **selective enhancement** are probably inappropriate.

- If selective enhancement is performed and properly declared in a submitted manuscript, then the editors and reviewers can decide if it is appropriate on a case-by-case basis.

## 1.7 Filters degrade data

Use of software filters to **improve image quality** is usually not recommended for biological images.

- Do not use filters from commercial packages that were developed for the print industry! They are usually undocumented and too aggressive.
- Selective filters can be used to pre-process the image for facilitating segmentation; do not just filter an image for the sake of it. If you do, provide proper arguments.

## 1.8 Cloning degrades data

Cloning or copying objects into a digital image, from other parts of the same image or from a different image, is very questionable.

- Users are often tempted to use tools like Photoshop's clone stamp tool to "clean up" a dirty preparation. **Do not do that!**
  - If the image is so bad to require retouching, you should probably acquire a new one.
- The use of cloning or copying techniques to specifically create objects in an image that did not exist there originally is research misconduct (falsification, fabrication).
  - Examples of misconduct would include copying gel bands into an existing gel image to create a new result, and/or any other image "seamlessly" created from the combination of portions of two or more images.
- In some instances the combination of two images into a single publication figure is appropriate.
  - The figure needs to make it obvious to the editor, reviewers, and journal readers that the two parts came from separate images.
  - A common example would be aligning lanes from two or more gels into one figure. Most journals require a line or small gap between the combined images to clearly show that they are from separate gels.

## 1.9 Making intensity measurements

Intensity measurements should be performed on uniformly processed image data, and the data should be calibrated to a known standard.

- Intensity measurements are difficult to perform in an appropriately rigorous and scientific manner.
  - Many variables must be considered and controlled for, before the data can be regarded as meaningful:
    - \* uniformity of sample preparation techniques;
    - \* optical aberrations and electronic noise to name a few;
    - \* intensity measurements would include the analysis of colocalization data, which is essentially the comparison of intensity data from two different wavelengths.
  - A particular difficulty with intensity measurements is the need for a robust and repeatable standard to calibrate or normalize the data. Unfortunately appropriate standards are not always available for purchase. Developing home-made standards is not a trivial task.
  - The many types of microscopes that are used to capture fluorescence images are subject to a number of known fluctuations over time, as well as other physics/electronics limitations.
- In general, intensity measurements should be performed on raw data to avoid potential artifacts.
  - If normalization, calibration, restoration, or any other image processing is performed on the data, it should be performed uniformly across all the data, and the procedures should be carefully described in the methods section of the publication.

## 1.10 Lossy Compression Degrades Data

Avoid the use of lossy compression.

- Image file compression comes in two basic types.

- Loss-less file compression reduces the size of the image file while maintaining the integrity of the image data. An example of loss-less file compression is the compressed Tiff file format, which uses the LZW algorithm. This type of file compression is appropriate for scientific data, although compressed Tiff files are not universally supported by imaging software.
- The most common form of lossy file compression is that provided by the Jpeg file format. Jpeg is an ISO/ITU standard file format for storing images that uses the discrete cosine transform to compress the file. The Jpeg file format is not suitable for scientific images because of the changes that the transform makes to the image data.
  - \* An excellent tutorial demonstrating the issues with using Jpeg for scientific images is available at Florida State University's Molecular Expressions website. See: <https://micro.magnet.fsu.edu/primer/java/digitalimaging/processing/jpegcompression/>.

## 1.11 Issues with Magnification

Scientific digital images are data that can be compromised by inappropriate manipulations.

- Digital images of real world objects sample the object such that each pixel in the image has a scale. This scale may be in meters per pixel for satellite images, or in tenths of microns per pixel for microscope images.
- Ideally, the scale is the same in both the X and Y dimensions; however, this is not always the case.
  - In confocal microscopy and other sectioning techniques, the XY pixel also represents a volume, because the image includes a Z dimension. The Z dimension in confocal microscopy is typically larger than that of the XY resolution, an issue that can lead to misinterpretation if not accounted for.
  - It is imperative that the scale of the pixels in XY and Z be maintained so that features in the image can be correctly interpreted.
- Microscopes can, in several imaging modes, visualize (but not necessarily resolve!) objects smaller than the diffraction-limited resolution of the instrument.
  - Sub-resolution objects typically appear to be the same size as objects that are at the actual diffraction-limited resolution; however, this is an artifact. Because of this problem, measurements of the size of objects that are at, or near, the diffraction limit are very suspect.
- Another important issue with sampling small objects using digital image capture is the need to correctly oversample the object.
  - The Nyquist sampling theorem suggests that a point object should be oversampled at least two times in X and Y. Because adequate contrast is essential to correctly resolve structures in microscopy, a 2.5-3 times oversampling is more appropriate.
    - \* Undersampling an image (using too few pixels to accurately describe a small feature) can yield artifacts that masquerade as real structures, which can lead to misinterpretation of the image data.
    - \* Oversampling does not lead to artifacts; however, it does not increase the diffraction-limited spatial resolution of the image and increases the data size.
- Scientists studying live cells in time should also consider oversampling on the temporal scale to avoid artifacts.

## 1.12 Issues with Pixels

Be careful when changing the size (in pixels) of a digital image.

- Because images may not fit the journal page while in their original size, users typically reduce or enlarge the image.
  - Changing the size of an image (the number of pixels in X and Y) can introduce aliasing artifacts.
  - Decreasing the image size in pixels reduces the spatial (XY) resolution in an image.
    - \* If the size reduction is not by a power of two, the software program performs an interpolation to create both a new XY resolution as well as the intensity values of each pixel.
    - \* If the image has regular, repeating structures, size reductions can create moiré artifacts.
  - Increasing the image size causes the software program to interpolate the new XY resolution and pixel-intensity values. Enlarging an image does not increase the spatial resolution; to the contrary, it may make specific features more difficult to resolve because aliasing artifacts tend to make the edges of features less distinct.

- In either case, one should insert a magnification scale bar prior to changing the total number of pixels in an image (magnification may be nearly impossible to calculate afterwards).
- One should resize an image only once. This prevents the resizing artifacts from adding up.