Flare identification with Convolutional Neural Net.

Preamble and Method selection

Lens flare occurs through internal refraction of an external bright light source, with a low angle relative to the lens normal: either a saturated white point in the image, or refracted 'ghosting' of different possible colours. A Convolutional Neural Net (CNN) is chosen for its simplicity, combined with some straight forward image filtering to achieve the desired outcome.

Algorithm Function.

The algorithm performs two operations. Image filtering stage (resizing, thresholding and smoothing) then processing; The training and input image sets are filtered identically.

Image filtering and preparation:
The input training images must be reduced and formatted for the CNN training stage. The images are resized (with the *cv2* library): a trade-off exists between image size and accuracy – larger images result in longer training time but return higher accuracy. Processing and training time goes approximately by 30 times the image size. A range of image sizes are tested to optimise the trade-off. As flaring can manifest as white (all image layers) or colours (some layers), the colour information is retained for training. A smoothing kernel (3x3) is applied to reduce complexity, and the images are thresholded at slightly less than the maximum value of 255: i.e. pixel values less than some value are zeroed - $P_{yxz} < 200 = 0$.

The convolutional Neural Net.
The CNN uses the *tensorflow* library to build a five layer convolutional net – with as many inputs to the net as there are image pixels, and with two outputs – '0' or '1' to reflect 'no flare present' and 'flare present'. The CNN is trained with the 80 provided input images, though it is expanded by rotating and reflecting the training set images four times, to yield a training set of 640 images. The training/testing fraction is set slightly higher than convention, at 70/30 %, due to the low number (and poor sampling of parameter space) of training images. Training time is a strong function of image size: an image size of 256 pixels takes approximately 13 hours, while 32 pixels is complete in less than a minute on a laptop PC.

Results summary
The important relevant metrics here are 1) CNN accuracy and 2) CNN speed. Following training, the entire sample dataset of 80 images is passed through the processing algorithm and timed.Not shown here tests varying the kernel and thresholding levels. Highlighted is a compromise of 99% accuracy. Tests are undertaken using a low-end PC laptop with 2 GHz processor with 16 GB RAM.

Image size (pixels)	Training time	Processing time (80 images)	Accuracy		
			False Pos (from 40)	False Neg (from 40)	Overall
32 x 32	10 sec	1.7 sec	~1	~3	95%
64 x 64	1.25 min	2.3 sec	~2	~0	97.5%
128 x 128	**15 min**	**8.7 sec**	**~0**	**~1**	**98.75%**
256 x 256	15 hrs approx	1:55 min	~0	~0	100%

Further work

Approaches to improve mitigation of lens flares exist both in hardware and software. Most simply; the nature of a lens flare is such that adding a physical hood to the cameras can almost eliminate flaring caused by sources outside the field of view. Additionally, while not presenting a problem in this case, it seems plausible that specular reflection from water surfaces can confuse a flare detection algorithm – specular reflection can be mitigated with appropriately rotated polarized filters on the cameras.

This algorithm was tested on a low-end laptop computer, certainly not optimized for image manipulation without any dedicated graphical processing capability. Dedicated GPUs can improve processing speed by orders of magnitude and will save cost as well as time.

In this case, the sample size of image data was quite small: typically a CNN is built using thousands of training images. Moreover, it was contaminated by rotated fields, clustering of a few images of unique situations (i.e. pipes) etc. Success of a CNN depends on appropriate sampling of the expected environments, without representation of unexpected environments – training images should conform to a uniform rotation (i.e. “sky” at the top) and include an unbiased sample from all possible environments.