Report on the CR investigation

Karlo Mrakovčić

March 1, 2022

1 Data

The Hyper Supreme-Cam (HSC) aboard the Subaru telescope provided the data. The science images in "HSC/raw/all" and templates from "HSC/runs/RC2/w 2022 04/DM-33402" were used to create two datasets. Visit 11710 was chosen, for no particular reason. After attempting to create dedicated templates for this investigation, that idea was abandoned, because it proved too time consuming. This way the most recent templates, which have already been validated, were used.

The type of pipeline used is the most significant difference between two datasets. The first was made with the standard HSC pipeline and default CR handling (finding + masking + interpolation on the science images), while the second was made without the CR interpolation. After image differencing in the second dataset, CRs should still be visible and be recognized as sources on the difference images. This way, the "noCR" dataset created with the default pipeline should have no CRs in the sources, and thus no "postage stamps" of the difference image with CRs, while the "CR" dataset should have all CRs as well as the real sources.

All "postage stamps" have size of 30x30 pixels around the center of the source.

2 "CR" dataset

The CR interpolation was turned off in this dataset by setting the keepCRs: true parameter. Pipeline detects cosmic rays and flags them with the CR flag inside the science image.mask object, but they remain inside the science image, and thus inside the difference image.

2.1 CR flags

The mask of the "lsst.afw.image.exposure" object must be examined to determine which pixel is part of the CR. This is accomplished by bitwise checking for the presence of the CR bit. To figure out which "postage stamp" has CR, the masks of those stamps are examined, and if one or more pixels have CR

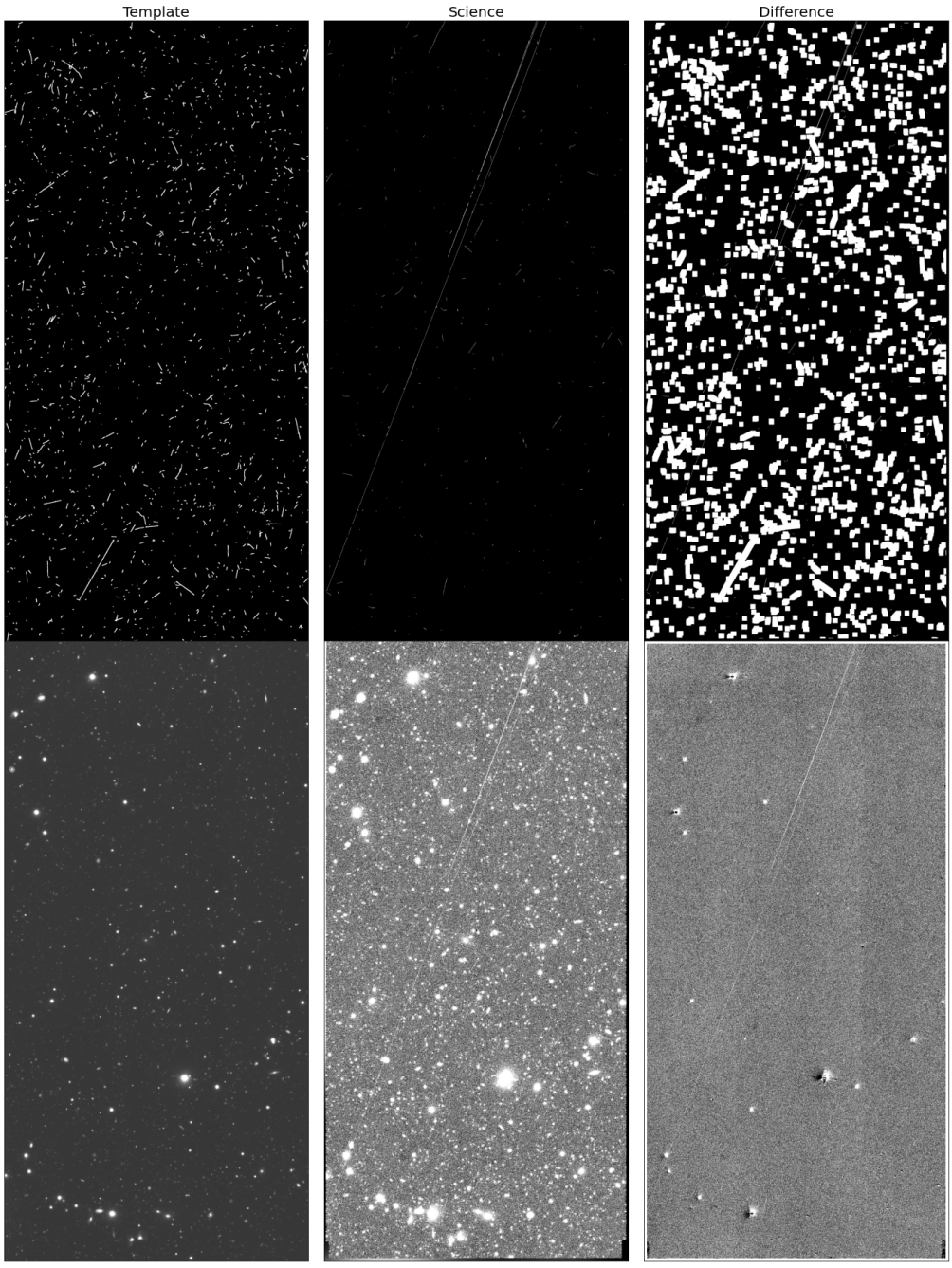


Figure 1: Top row showing CR flags while the bottom row showing *tanh* (exposures) of template, science, and difference image. Visit 11710, detector 42. It is clear that the difference mask has many more CR flags than science mask, and the template mask is not empty.

flags, the entire stamp is labeled as CR. It turns out that the science mask must be examined rather than the mask of the difference image postage stamps, because difference masks have a lot of "false positive" CR flags. Reason for that is that the CR flag inside difference images is also inherited from template, which should be empty, but it is not. The figure (1) shows the distribution of CR flags on the template, science, and difference exposures, as well as corresponding *tanh* value of the pixel intensity. *tanh* of the intensity was chosen as a quick means of coping with the extreme range that pixels have, from very dark sources, to very bright sources.

It is also worth noting that something what appears to be satellite trails are marked as CRs.

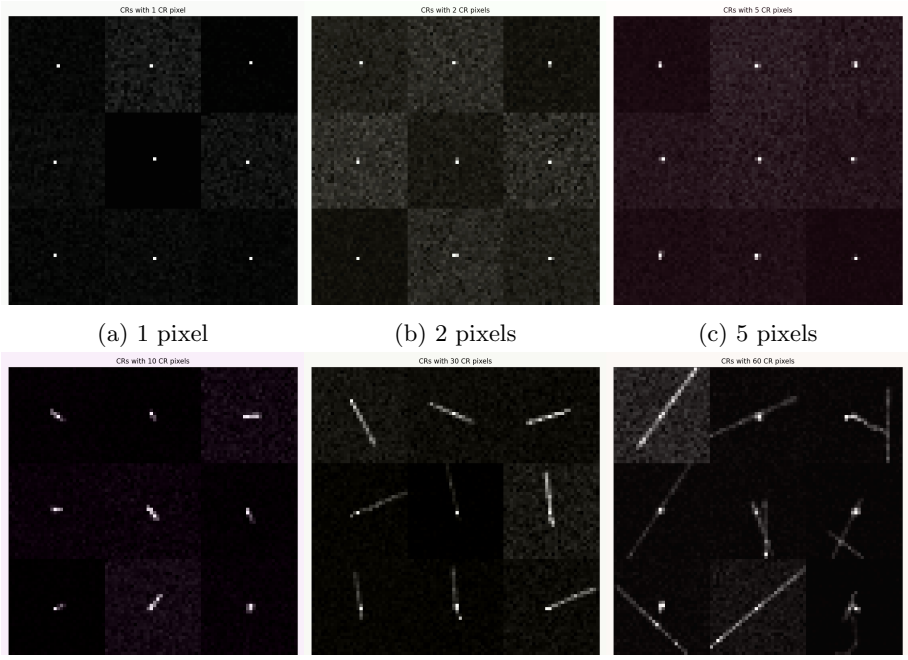
2.2 Cutoff

To fully understand how to label stamps one must understand how many CR flags usually are in the "CR" stamp. After understanding that some number of 1 pixel CRs exist cuttof was made that only one CR flag inside the stamp is enough to make the whole stamp labeled "CR". Figure (2) shows 1 pixel stamps, while figure (4) shows number of CR pixels inside CR stamps.

Figure 2: All of the examples of stamps with 1 pixel flaged as CR for visit 11710. It is clear that most of the stamps are in fact CRs so the cutoff=1 is chosen.

2.3 Object detector with CRs

It's interesting to see how many of the CRs that are still visible in the difference image are recognized as objects. Figure (5) depicts one of the visit's detectors.



(a) 1 pixel (b) 2 pixels (c) 5 pixels

(d) 10 pixels (e) 30 pixels (f) 60 pixels

Figure 3: Examples of stamps with different numbers of CR flags in its mask.

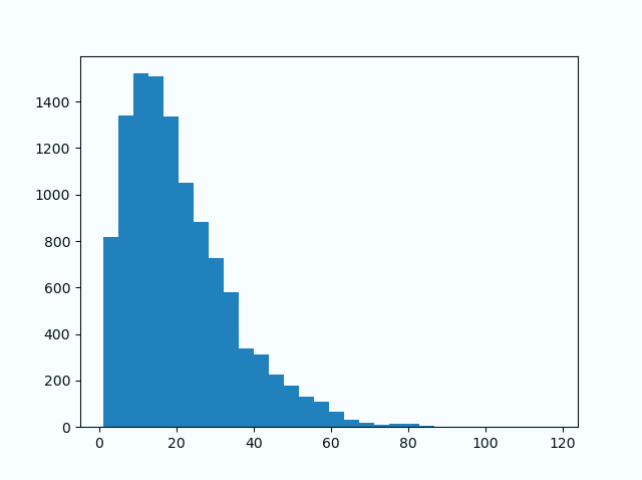


Figure 4: Histogram of number of CR flaged pixels per stamp. Each stamp has size of 30x30=900 pixels. Shown are only stamps with CR flaged pixels ¿0.

Only the pixels that are recognized as CRs are colored in white, while the red

squares resemble the "postage stamps" around the sources with CR label. The CR label was deduced by looking at the science mask on the stamp position, and marking it as a CR if it has more than 1 pixel marked as CR. The conclusion is that the majority of CRs are not recognized as objects, and the object detector is ineffective at detecting CRs as objects. Out of the total 878807 pixels with CR flags in science mask, only 232247 are inside stamps.

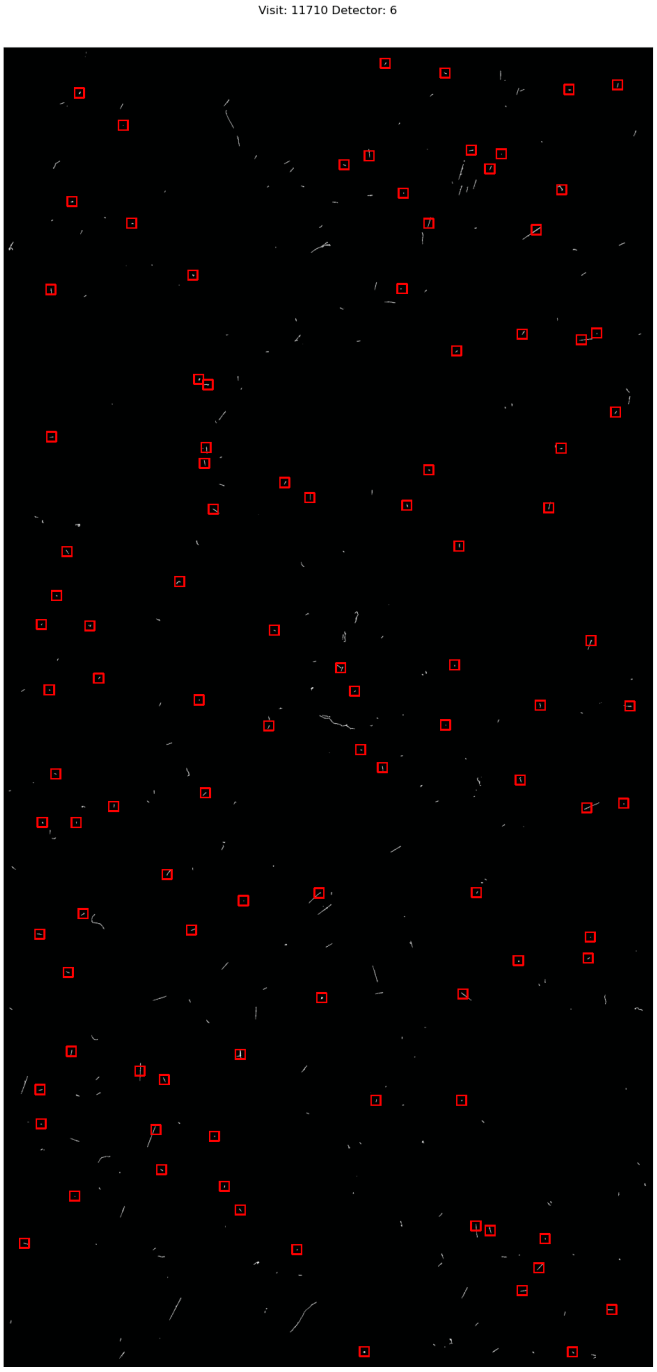


Figure 5: White pixels showing CR flags, while red squares show stamps labeled as CRs. Most of the CRs are not detected as a sources, especially longer ones.

2.4 CR stamps

After labeling and cutting out, there are a total of 18343 stamps, 11211 which are labeled as CRs and 7132 which are not. Figure (7) is showing the stamps labeled as CRs while the figure (6) are showing all the non CR stamps. It is clear that not all of the CR made its way inside CR labeled ones, and the doubious ones are shown in figure (8a). It is important to note that algorithm did not mislabel any source as CR, most mislabels are due to CR passing over the source.

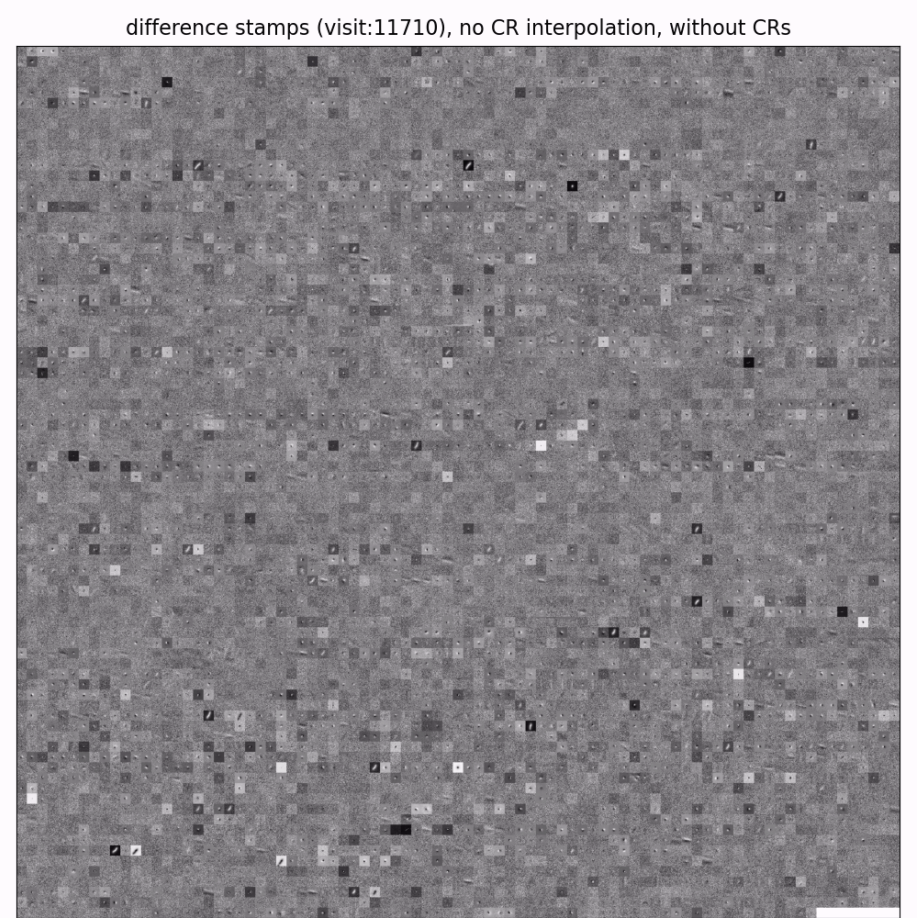


Figure 6: All postage stamps labeled as not CR. There are multiple suspicious stamps that look like stamps with CRs

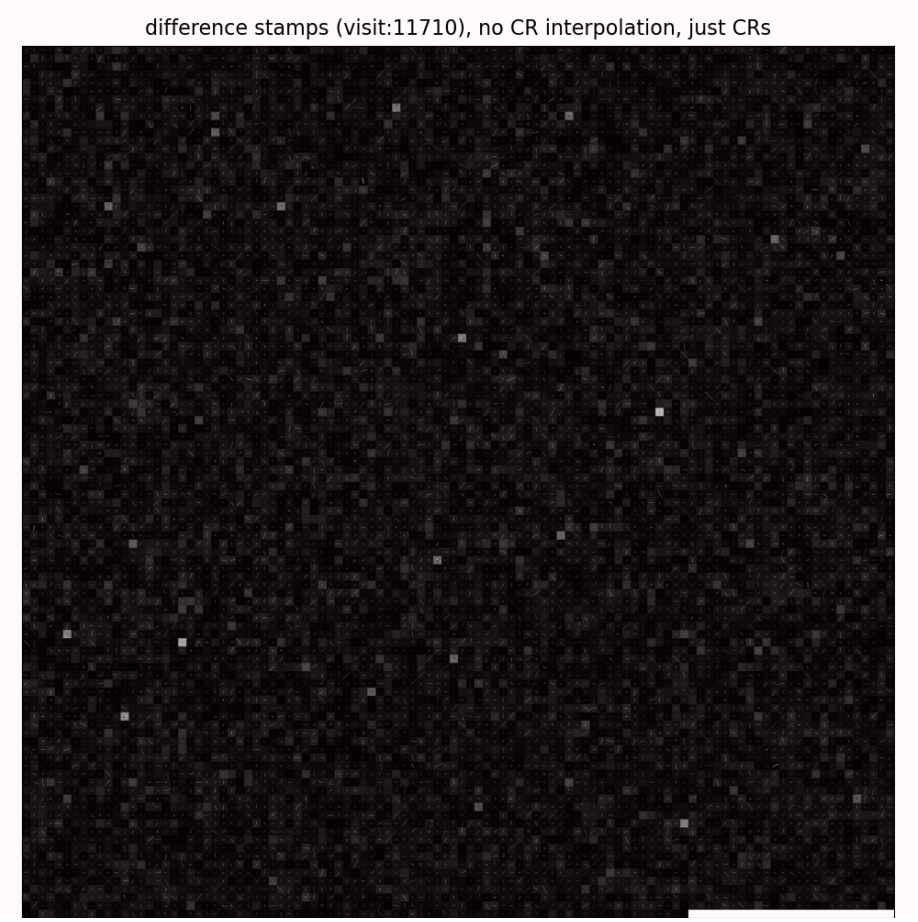
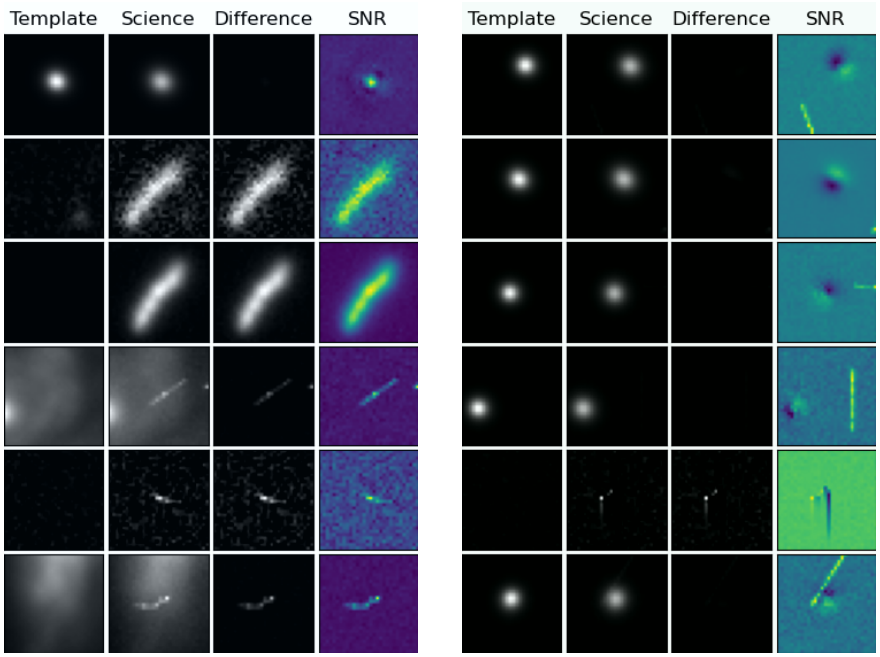


Figure 7: All postage stamps labeled as CR. Most of the "false positives" are cases where CR overlaps with a source.



(a) not CR (b) CR

Figure 8: Suspicious stamps that look like mislabels.

3 “not CR” dataset

This datased should in theory contain no cosmic rays. This however is not true, and similarly to figure (6), figure (9) contains some amount of CR candidates. Curiously some of the stamps still contain CR flags, at the places where the CRs happen at the same place where the object is, or where interpolation is faulty. This can be seen on figure (10). Out of 7419 stamps in this dataset, 227 contain CR flag, and 7192 do not. It is curious to compare that ”CR” dataset contains 7132 not CR flags.

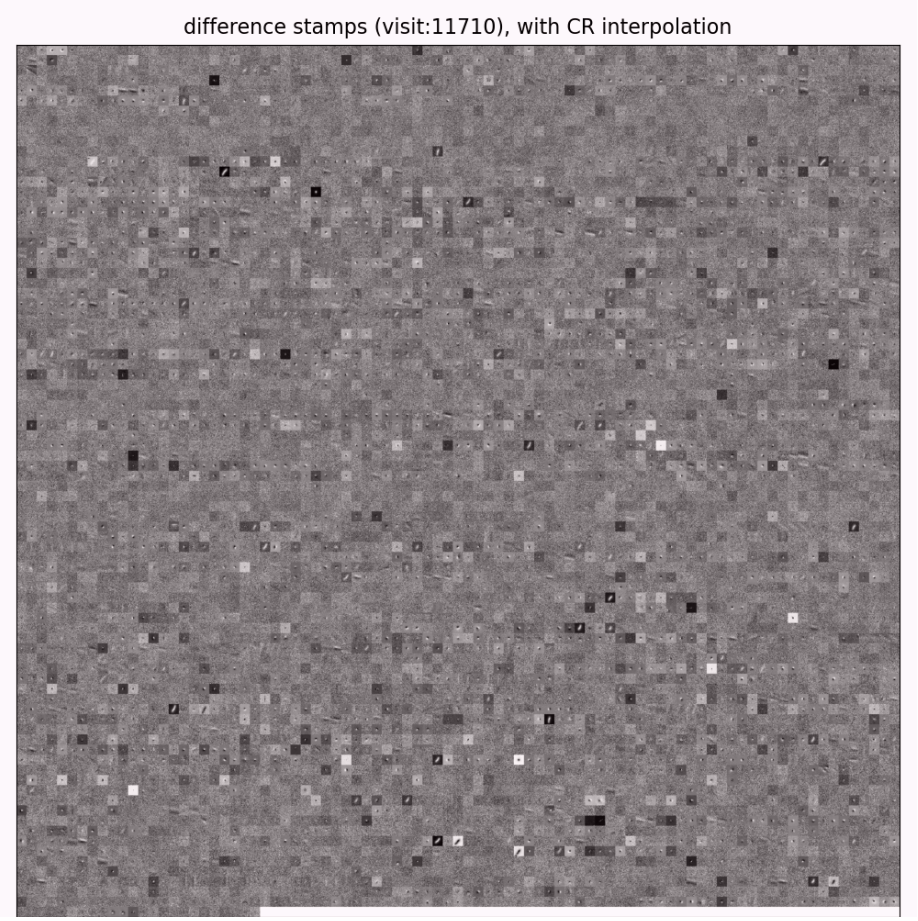


Figure 9: All stamps in not CR dataset. Most of the stamps are already in figure 6.

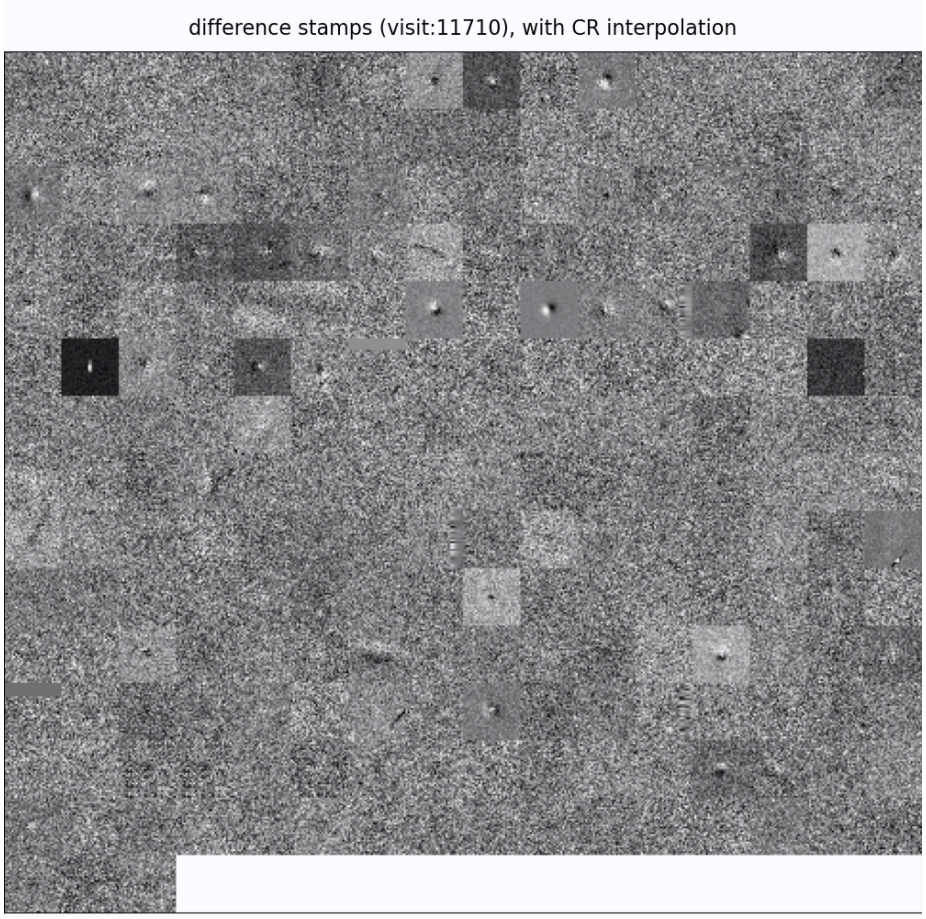


Figure 10: Only CR labeled stamps in CR dataset. Many stamps have artefacts of poor interpolation.