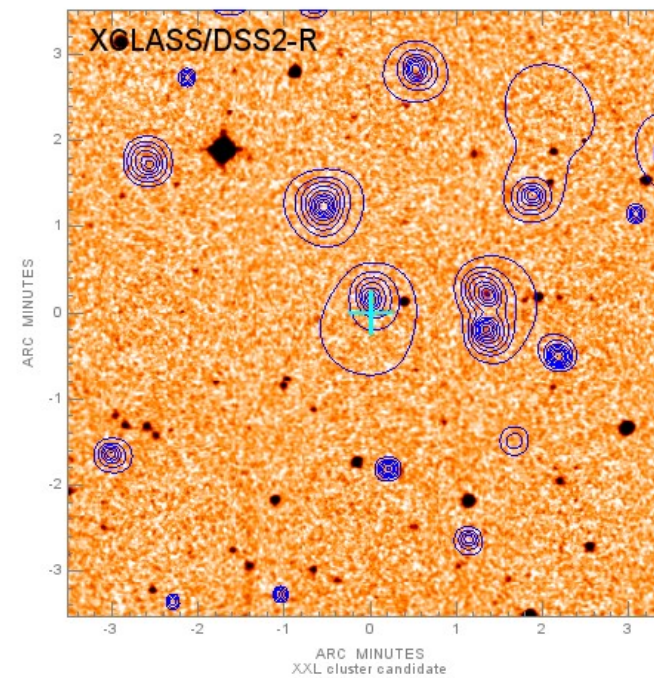
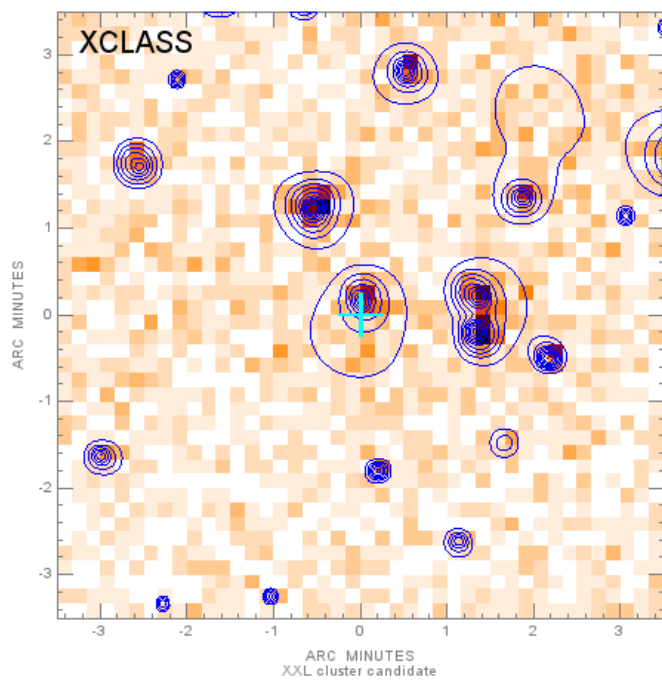
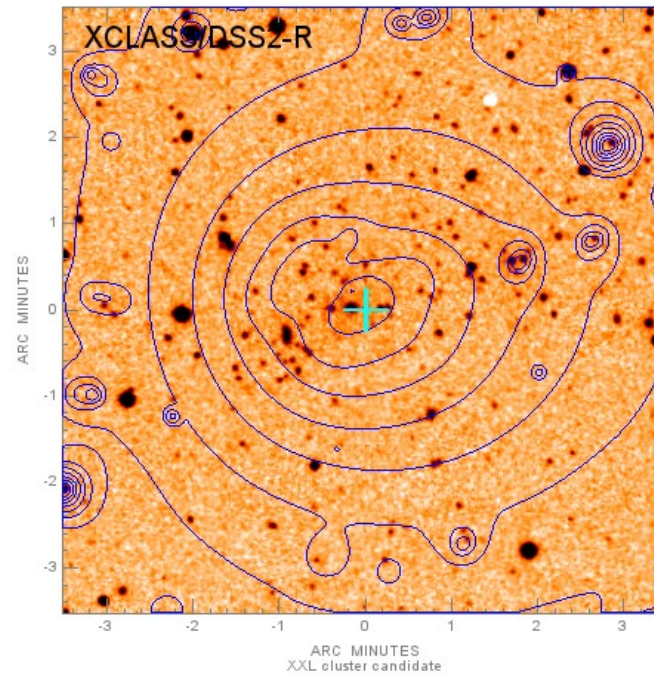
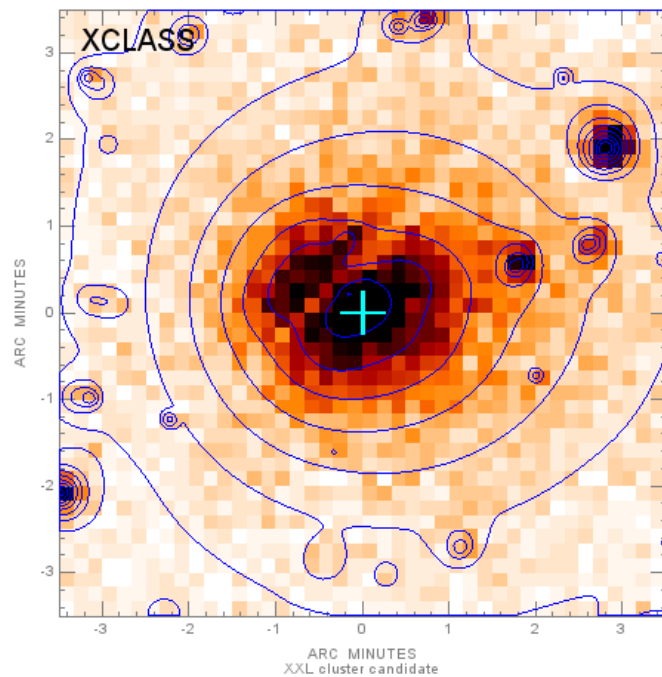
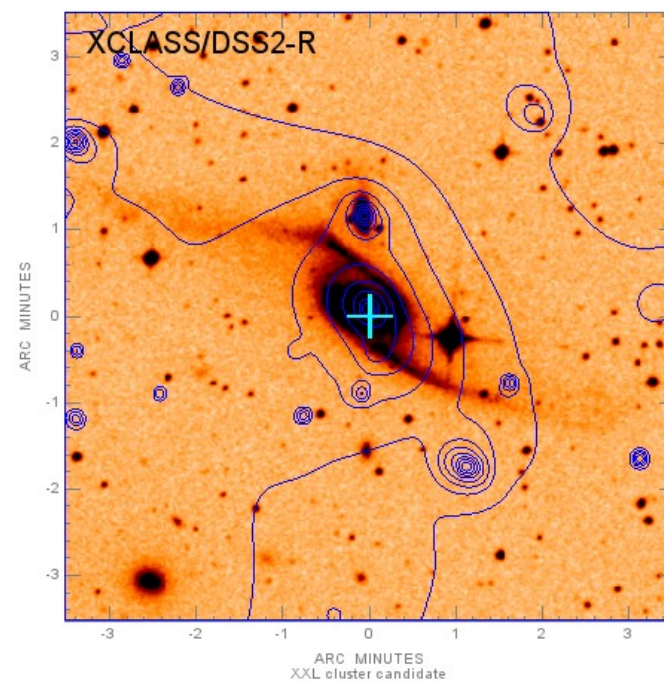
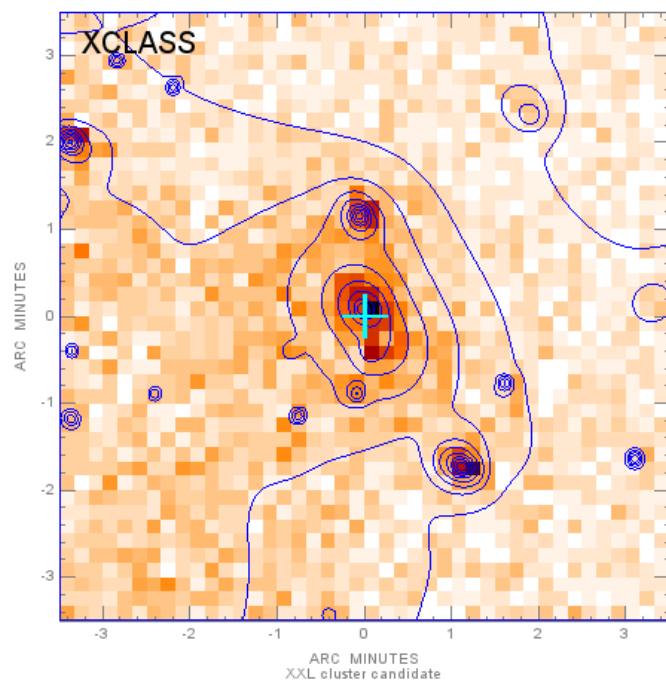
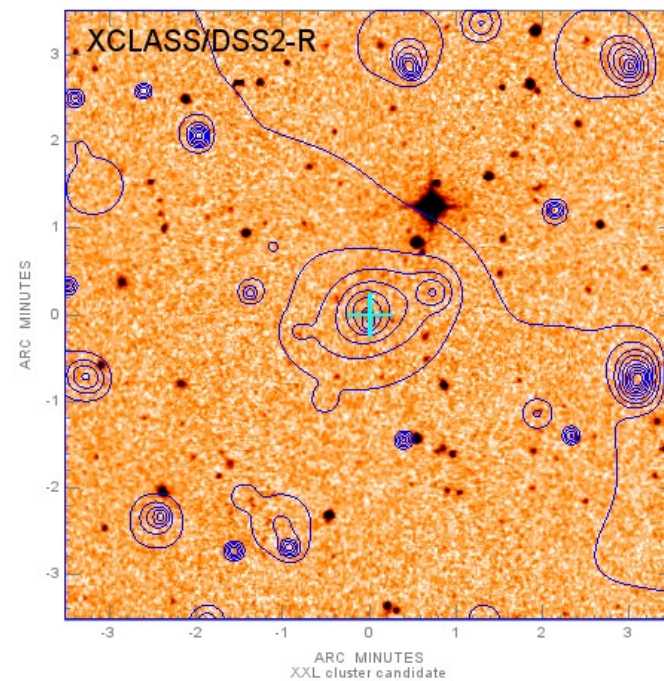
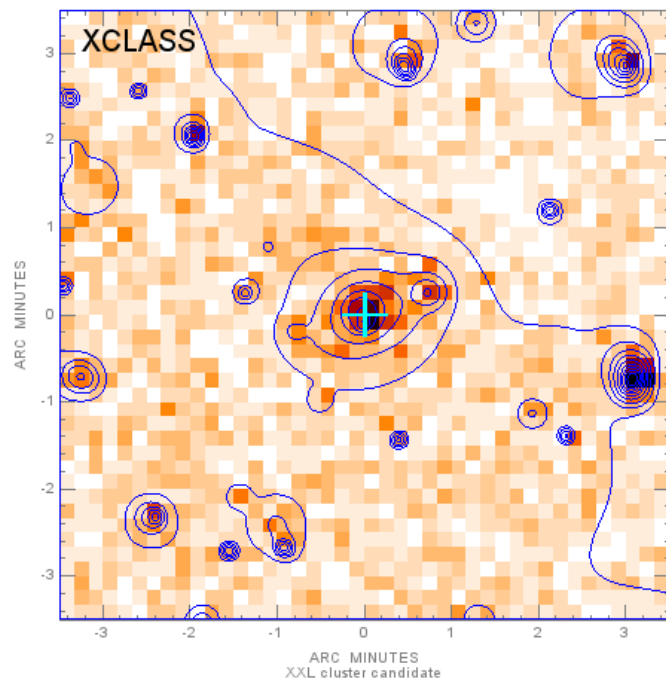


Multiwavelength classification of X-ray selected galaxy cluster candidates using convolutional neural networks

Matej Kosiba







The Hunt for Galaxy Clusters

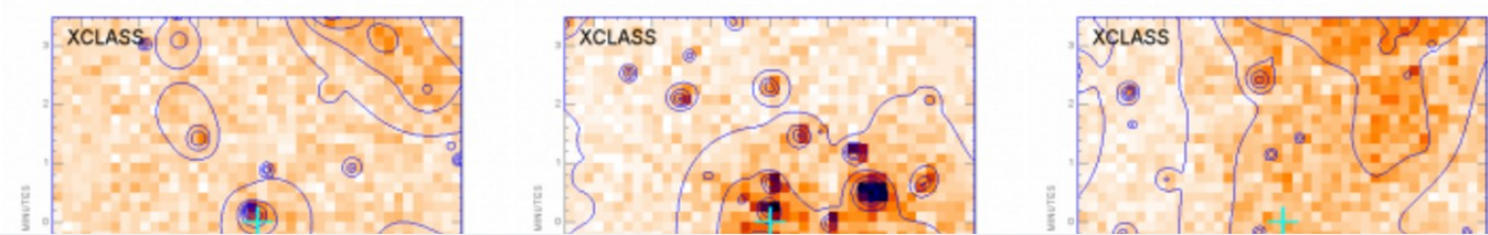
ABOUT CLASSIFY TALK COLLECT

Great work! Looks like this project is out of data at the moment!
[See the results](#) or [dismiss this message](#)

Explore the depths of the Universe
with Galaxy clusters

[Learn more](#)

[Get started](#)



No one is talking about **The Hunt for Galaxy Clusters** right now.

[Join in](#)

THE HUNT FOR GALAXY CLUSTERS STATISTICS

100% Complete

1 240

Volunteers

48 225

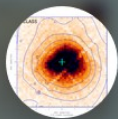
Classifications

1 600

Subjects

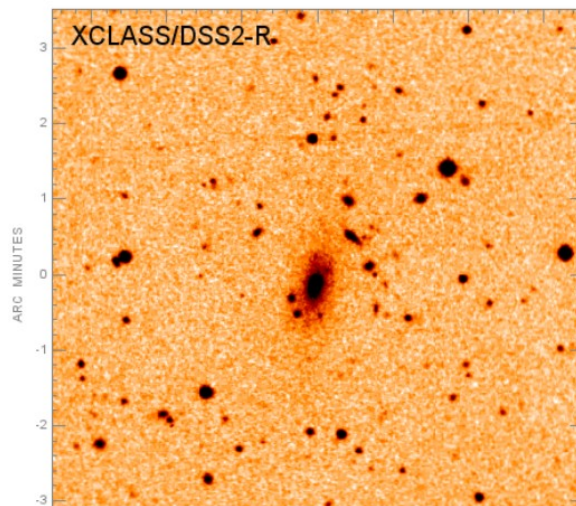
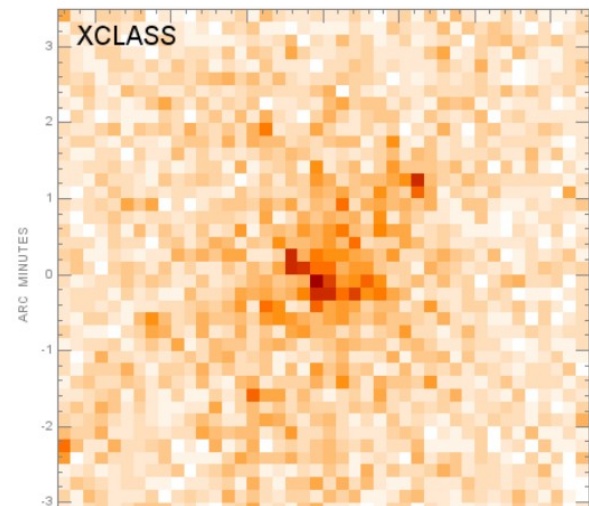
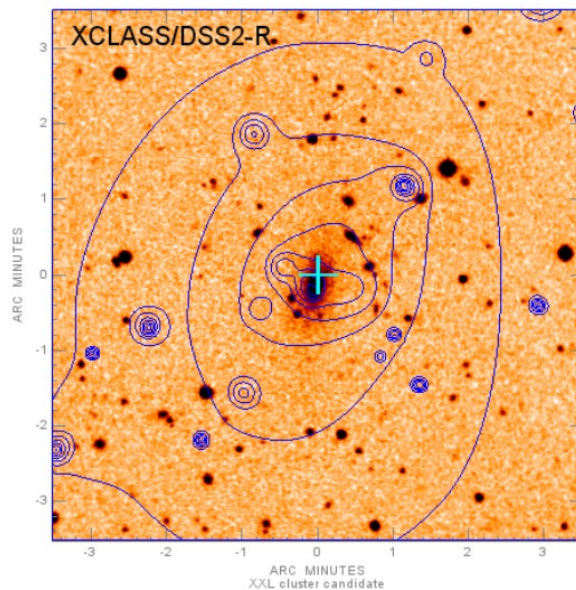
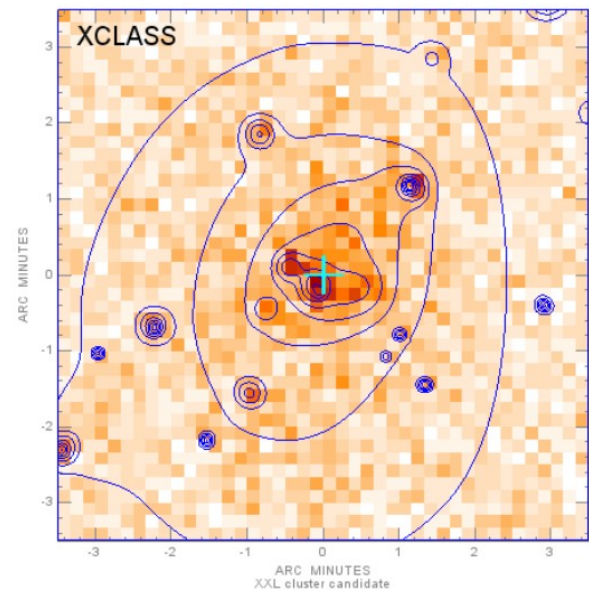
1 600

Completed Subjects



The Hunt for Galaxy Clusters

Give us your feedback via this short Google form <https://goo.gl/forms/KzThMUVW8E3AeSHH3>



TASK

TUTORIAL

Look at images 1 & 3. Do you see X-ray emission close to the centre of the image?

Emission close to the centre

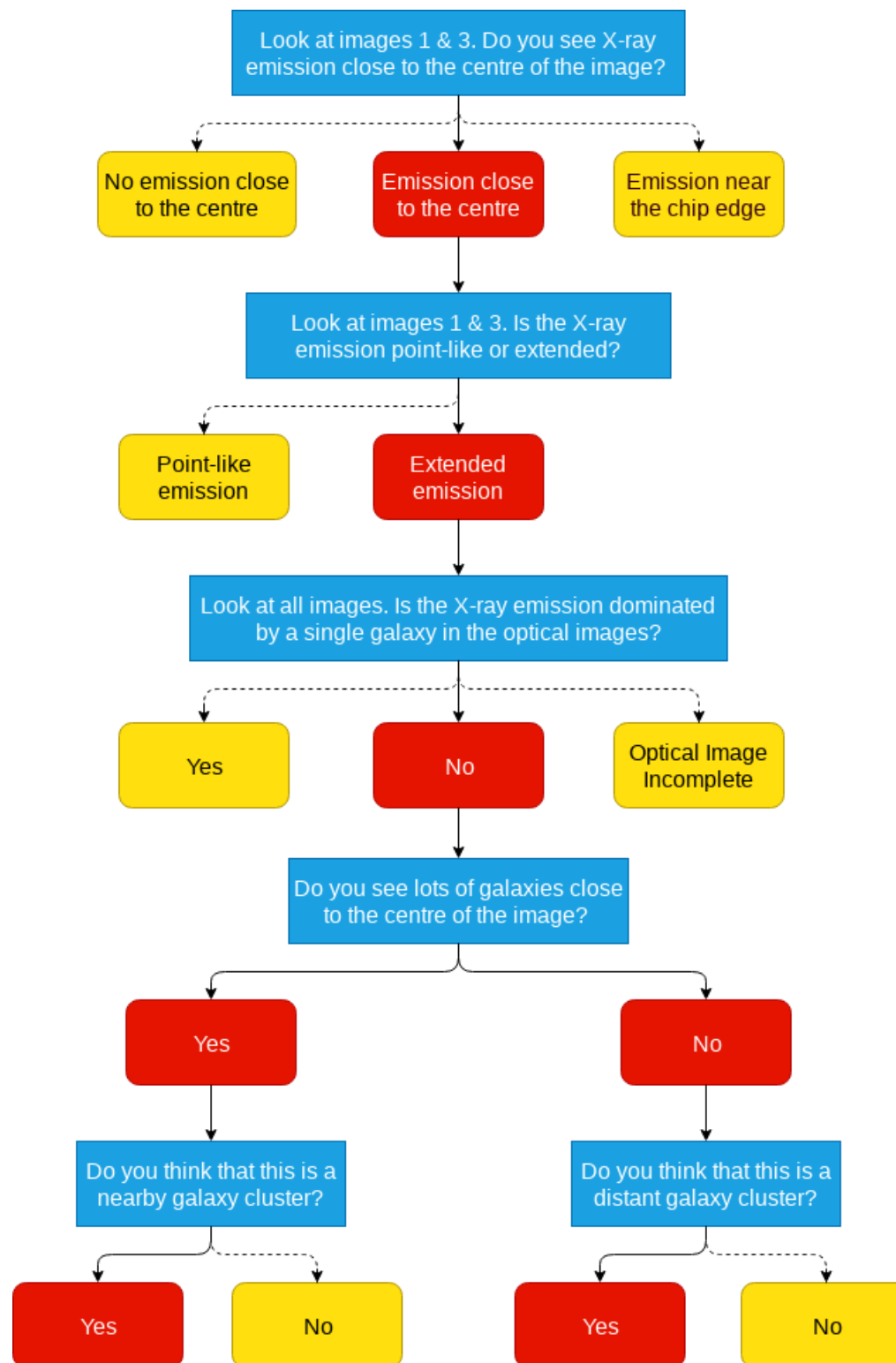
No emission close to the centre

Emission near the chip edge

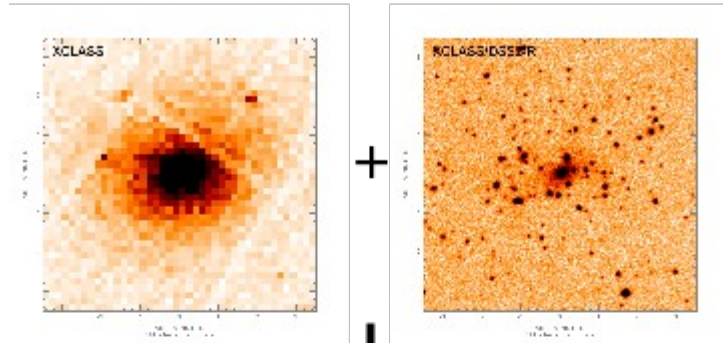
NEED SOME HELP WITH THIS TASK?

Done

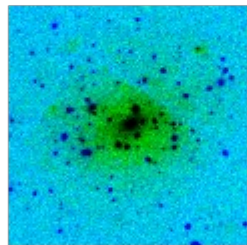




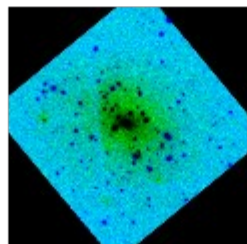
Input



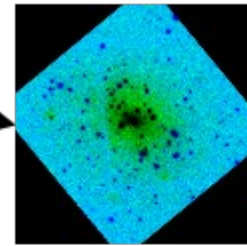
Preprocessing



Augmentation



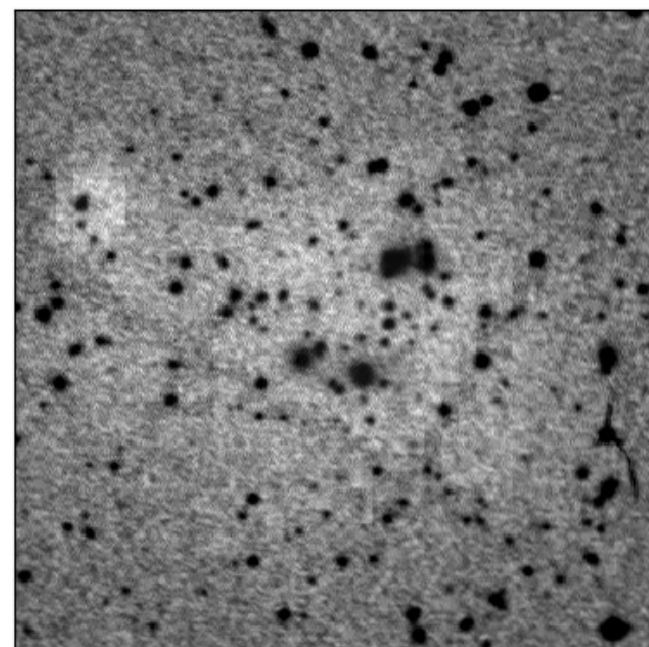
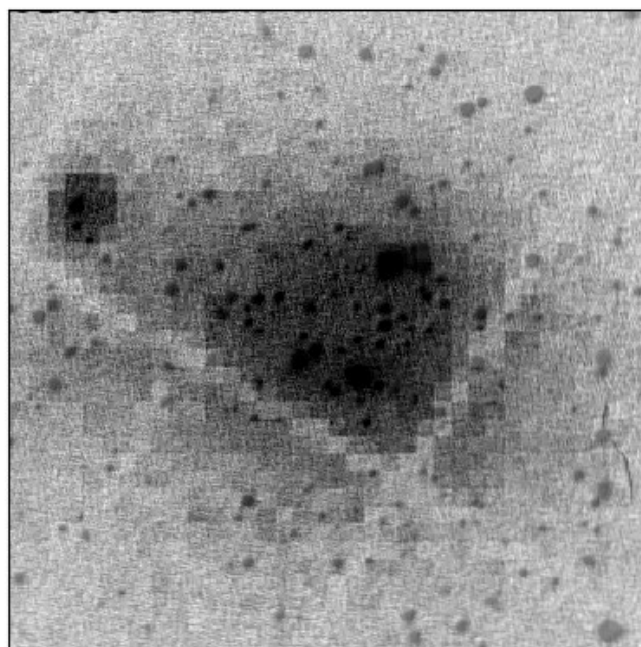
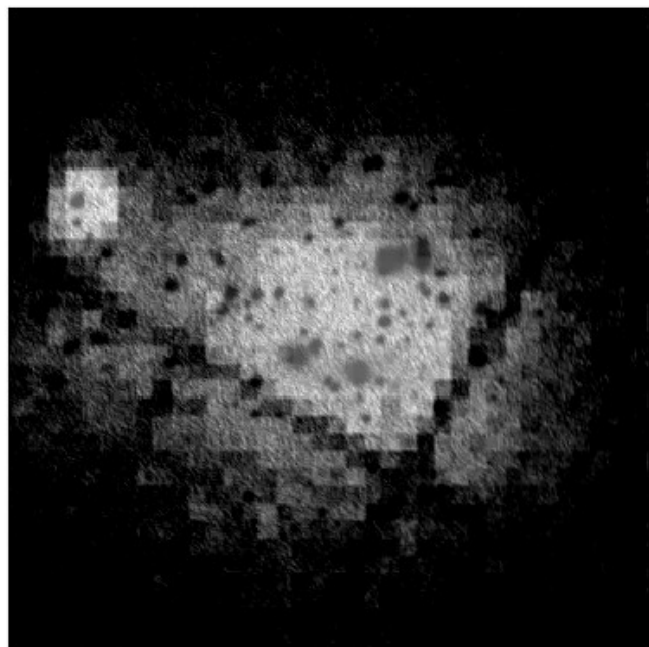
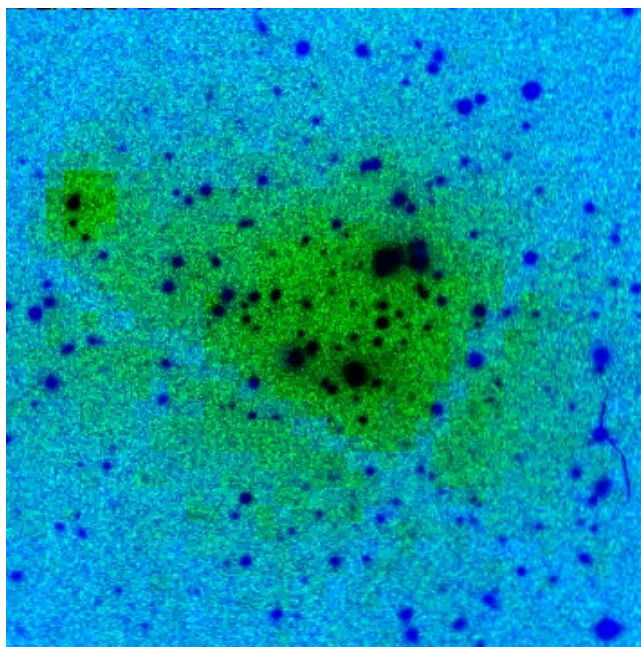
Augmented input

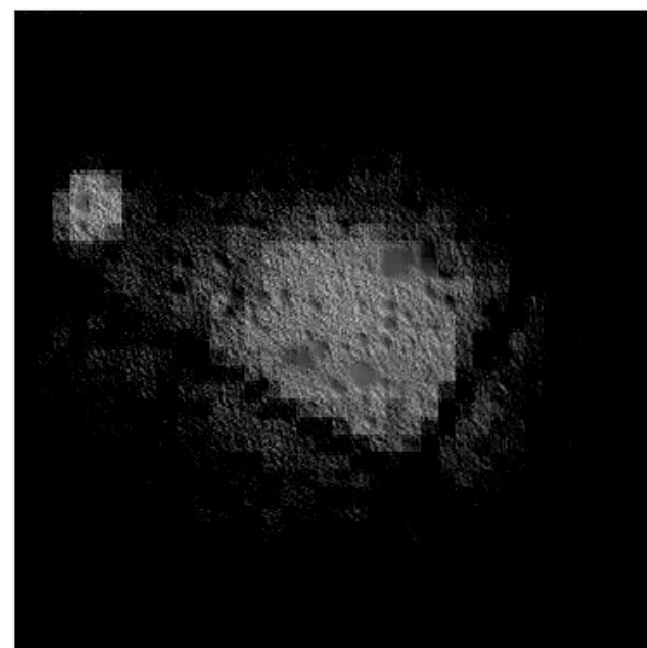
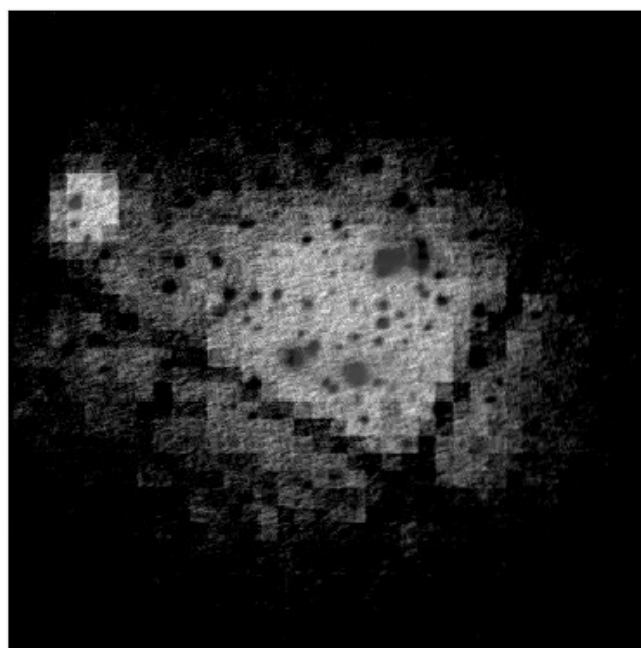
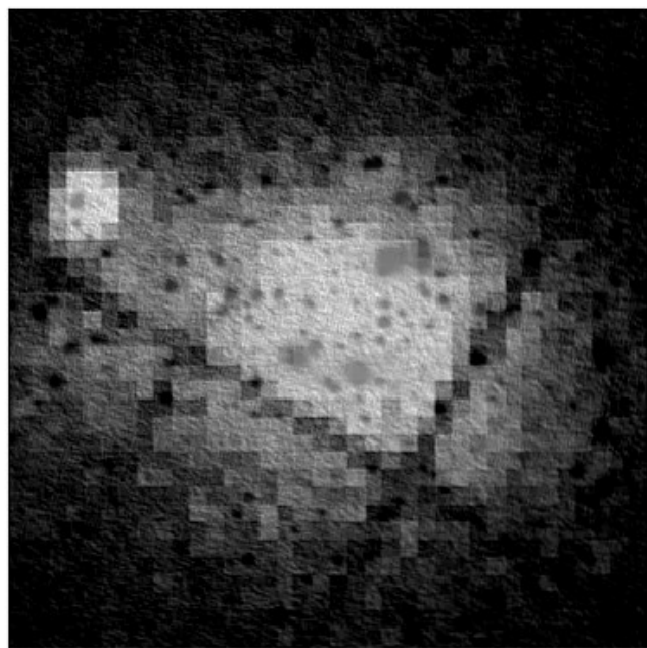
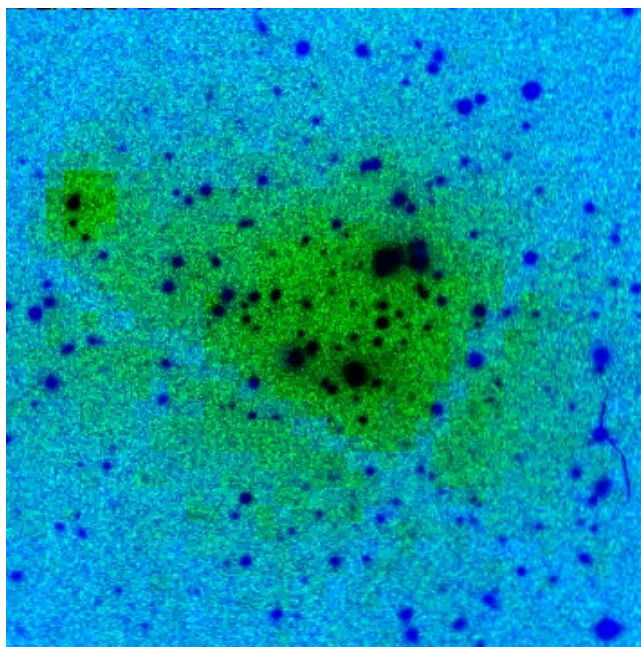


+ Label

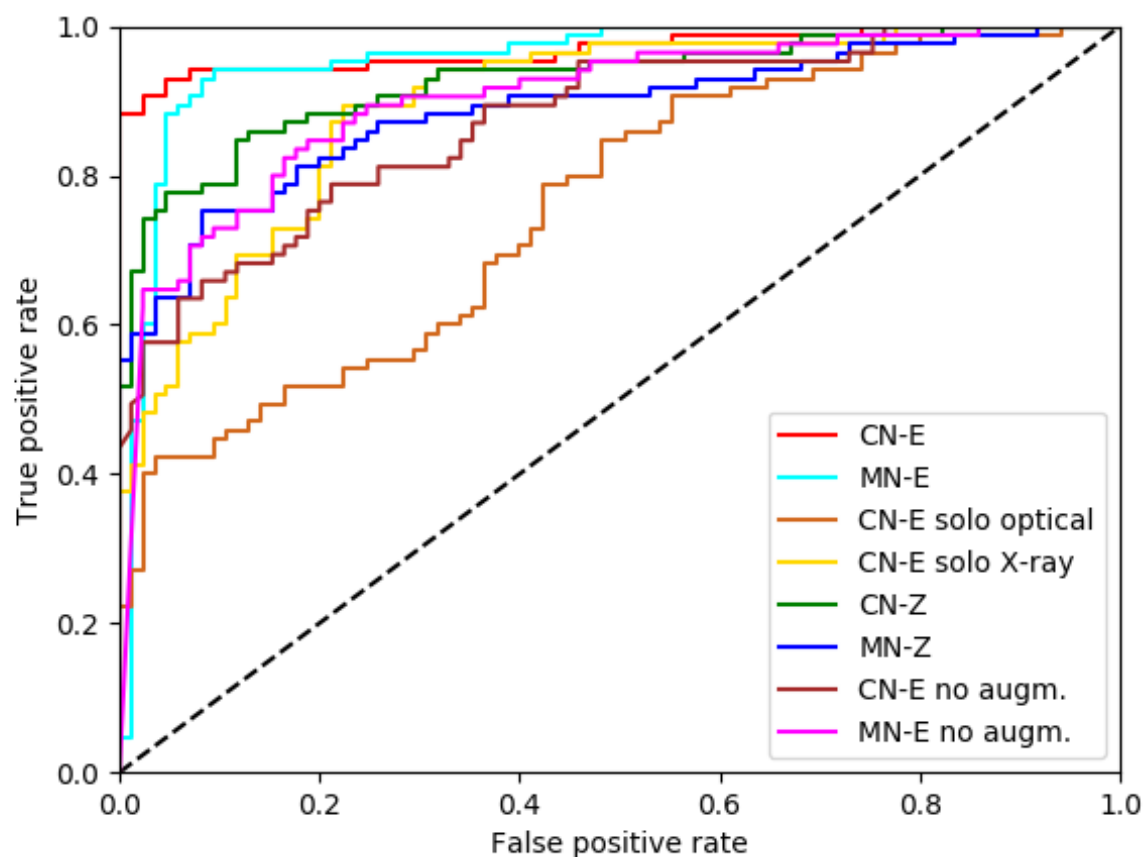
Convolutional neural
network

Predictions





network	A \pm std	P \pm std	R \pm std	AUC \pm std
CN-E	0.90 ± 0.03	0.89 ± 0.05	0.91 ± 0.03	0.96 ± 0.01
MN-E	0.88 ± 0.02	0.87 ± 0.03	0.91 ± 0.03	0.94 ± 0.01
CN-E solo optical	0.68 ± 0.02	0.64 ± 0.02	0.85 ± 0.04	0.77 ± 0.02
CN-E solo x-ray	0.81 ± 0.01	0.78 ± 0.03	0.86 ± 0.04	0.89 ± 0.01
CN-Z	0.82 ± 0.01	0.96 ± 0.01	0.67 ± 0.02	0.91 ± 0.01
MN-Z	0.79 ± 0.02	0.96 ± 0.03	0.62 ± 0.03	0.86 ± 0.02
CN-E no augm.	0.75 ± 0.02	0.70 ± 0.02	0.87 ± 0.03	0.87 ± 0.01
MN-E no augm.	0.81 ± 0.01	0.75 ± 0.02	0.91 ± 0.01	0.90 ± 0.02



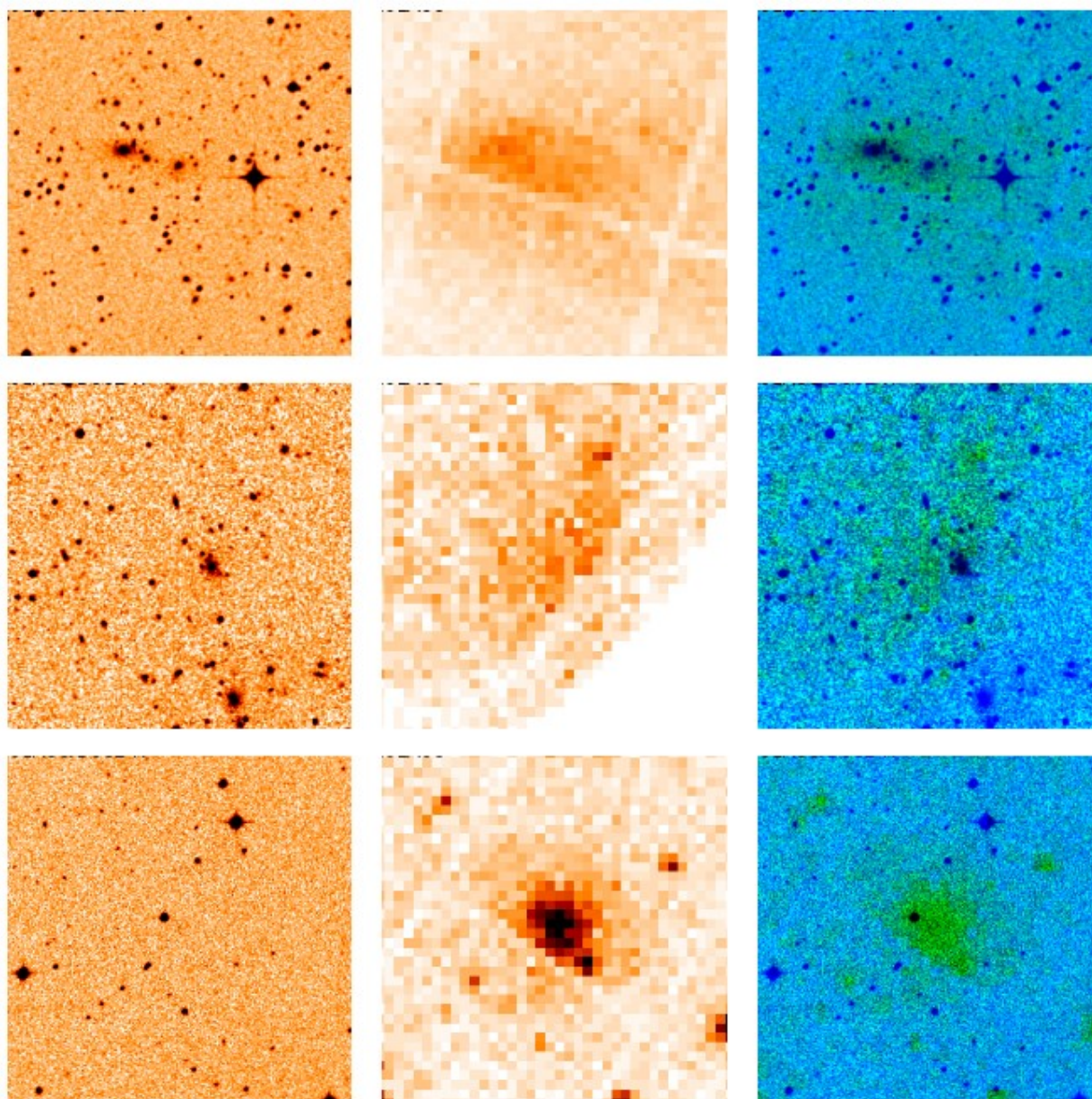


Figure 9. Spectroscopically confirmed galaxy clusters correctly classified by our custom network randomly selected from the test sample (TP). *Left:* optical, *middle:* X-ray, *right:* combined.

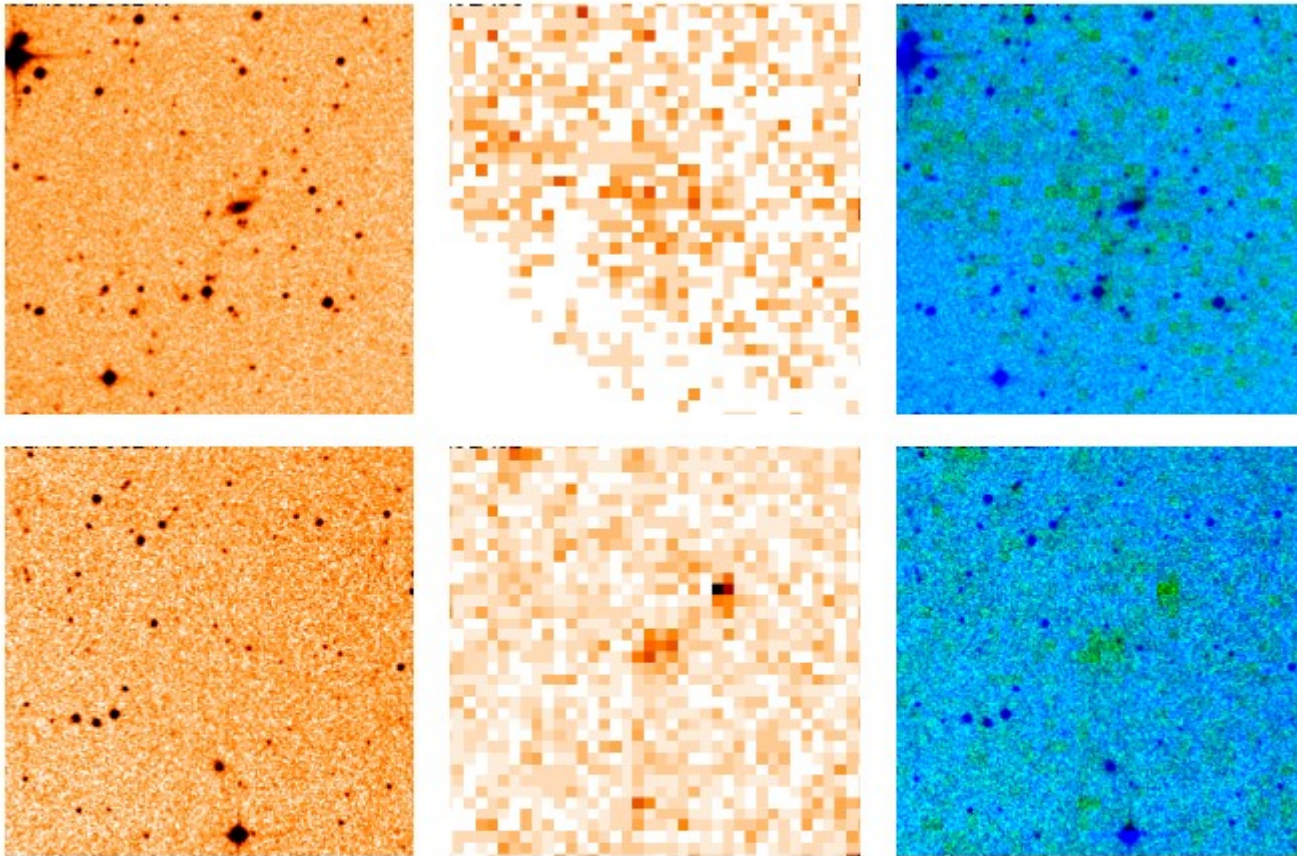


Figure 10. Non-galaxy clusters incorrectly classified as galaxy clusters (FP) by our custom network. *Left:* optical, *middle:* X-ray, *right:* combined.

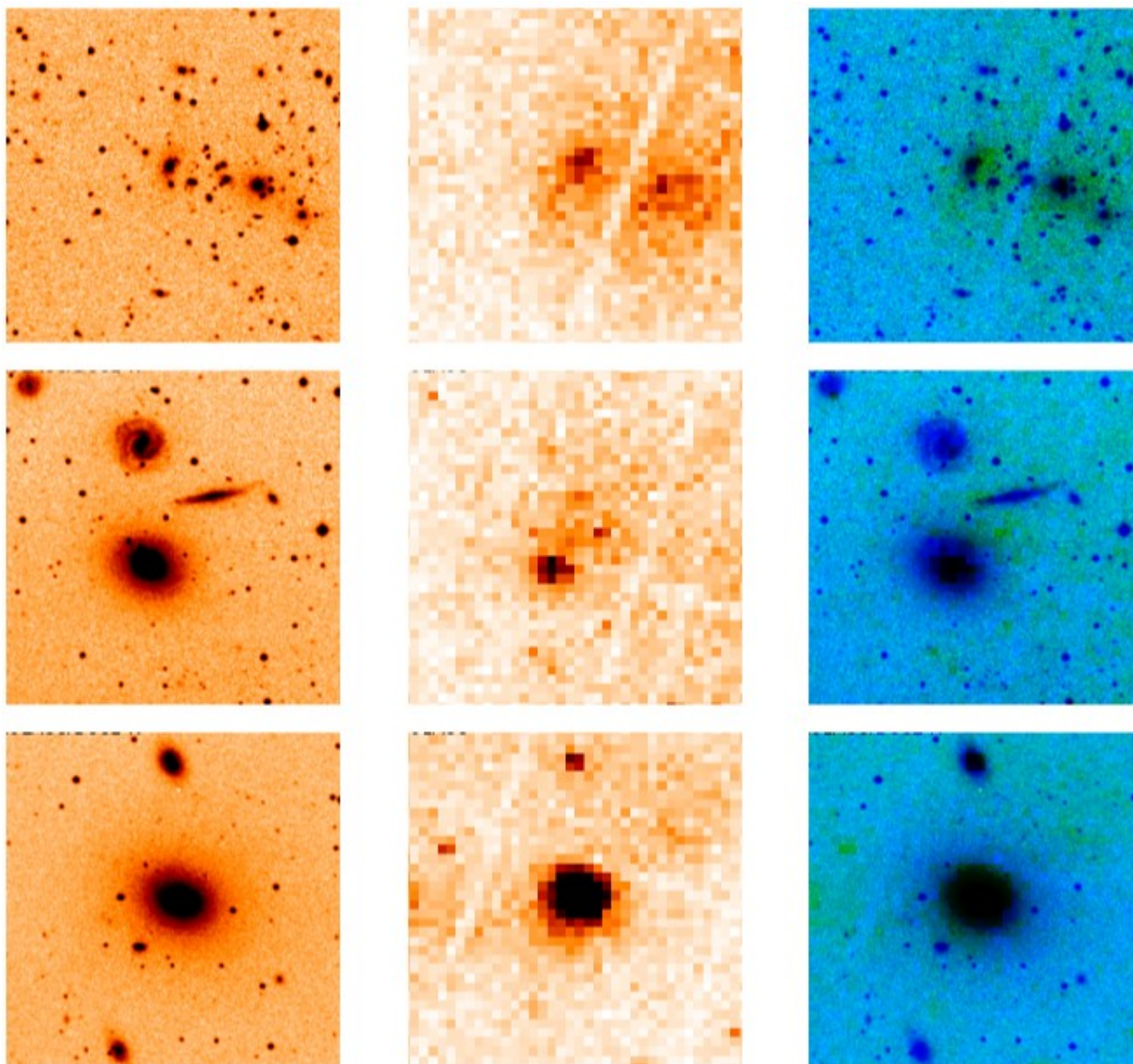


Figure 11. Galaxy clusters incorrectly classified as non-galaxy clusters (FN) by our custom network. *Left:* optical, *middle:* X-ray, *right:* combined.

Conclusion

- 62% agreement between Zooniverse volunteers and experts, however an extremely pure sample of clusters (99% agreement).
- Our hand-made convolutional neural network achieved the best average accuracy, 90%.
- Future plans:
 - Galaxy cluster catalogue.
 - Object detection approach.
 - Utilisation of other wavelengths, radio, infrared and/or SZ effect.
 - Conducting cosmological studies.

Bonus Slides

Table 1. The results of cluster classification by Zooniverse volunteers on two data sets, 404 objects are those classified by both, scientists and Zooniverse volunteers, the 170 objects data set is a subsample of the 404 objects, where 85 objects are spectroscopically confirmed clusters and 85 are objects classified as non-clusters by experts.

Data set	Zooniverse Classifications	TP	TN	FP	FN	accuracy	precision	recall
404 objects	unweighted	69	150	0	185	0.542	1.000	0.272
404 objects	weighted	102	149	1	152	0.621	0.990	0.401
170 objects	weighted	55	84	1	30	0.818	0.982	0.647

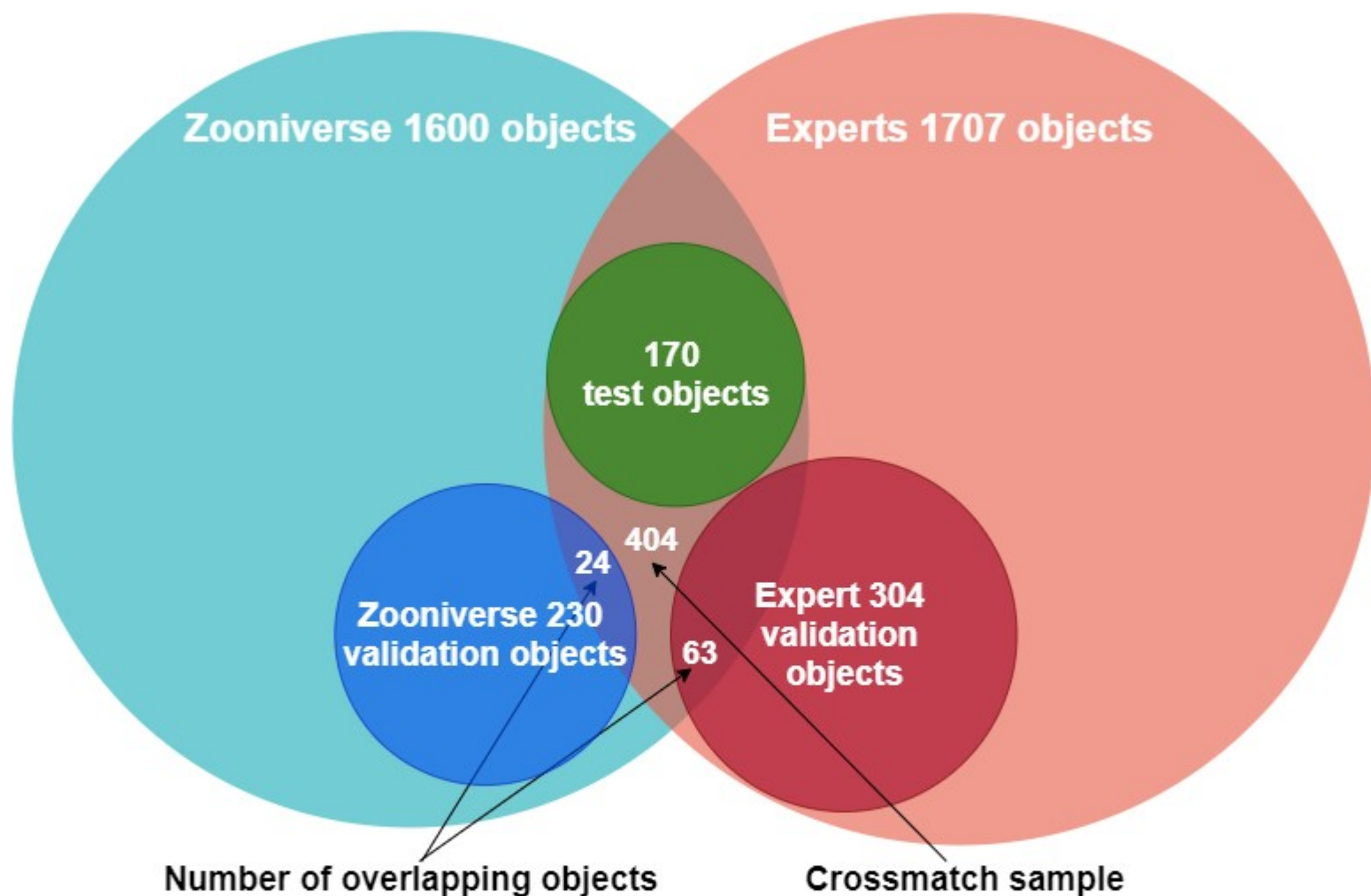
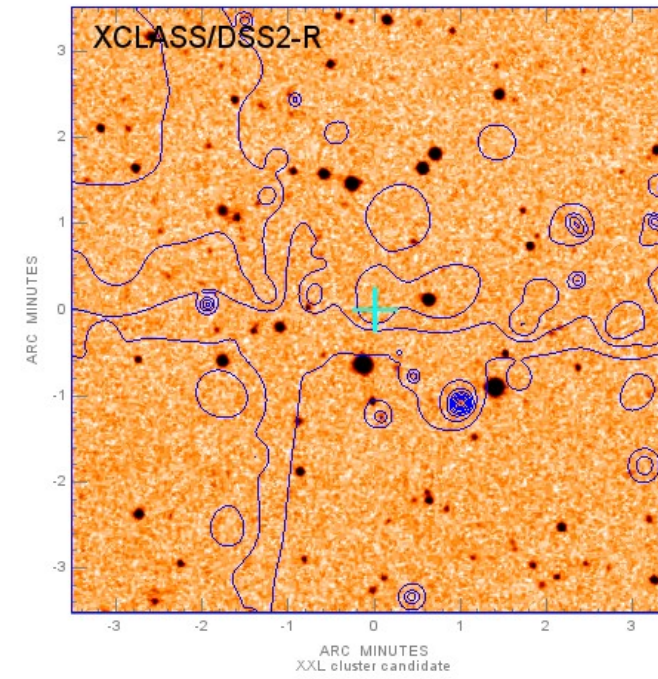
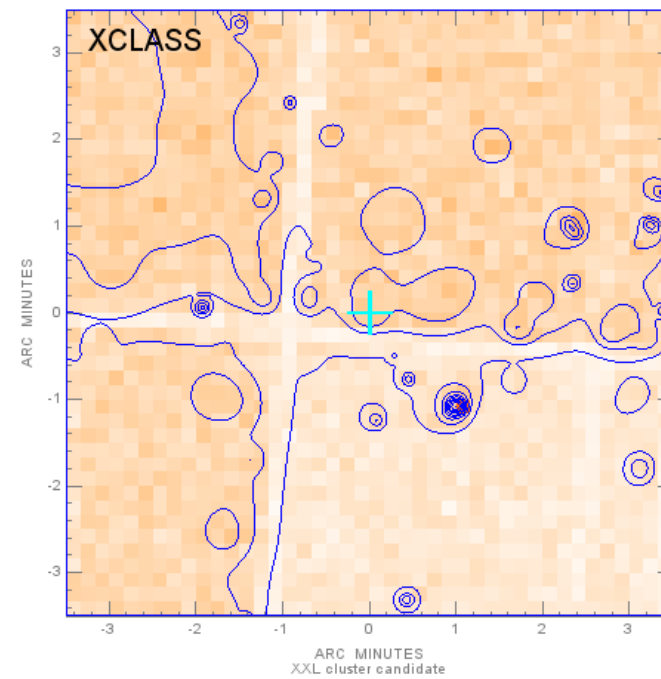
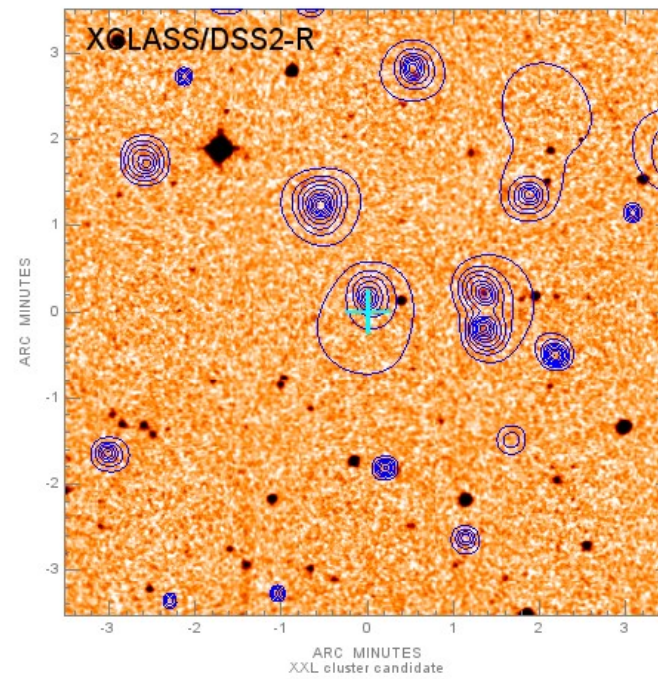
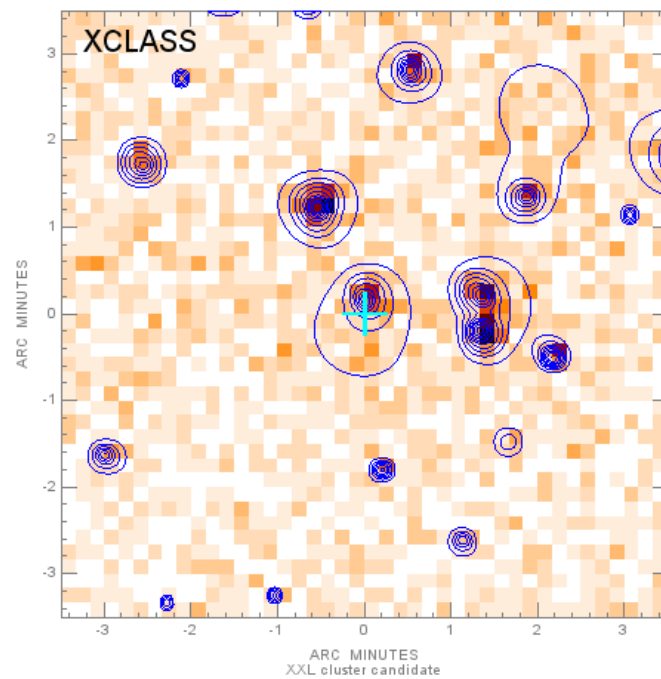


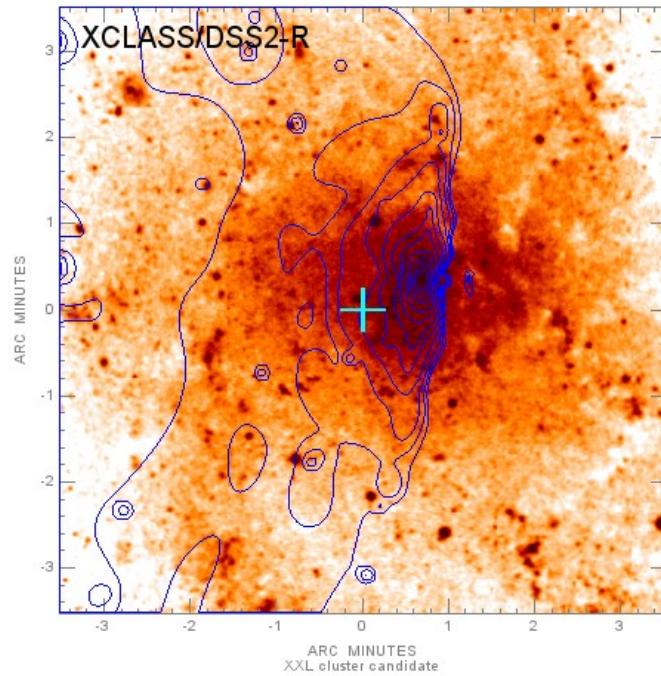
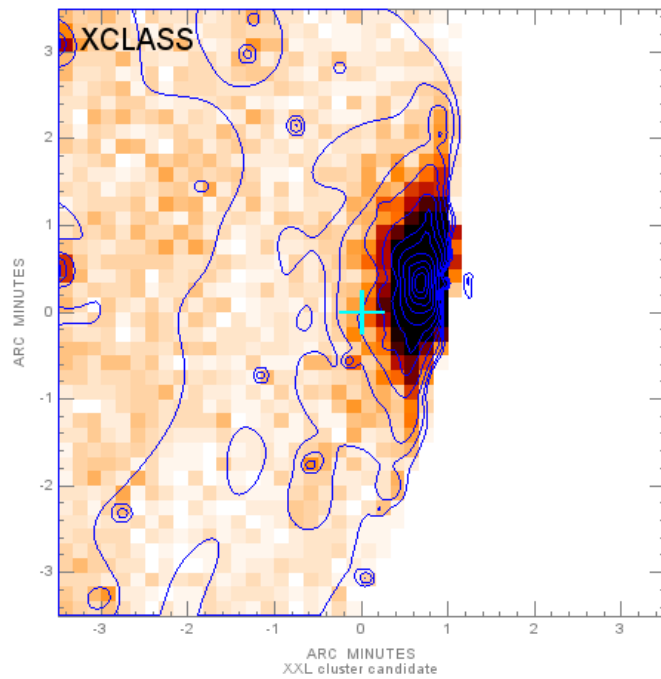
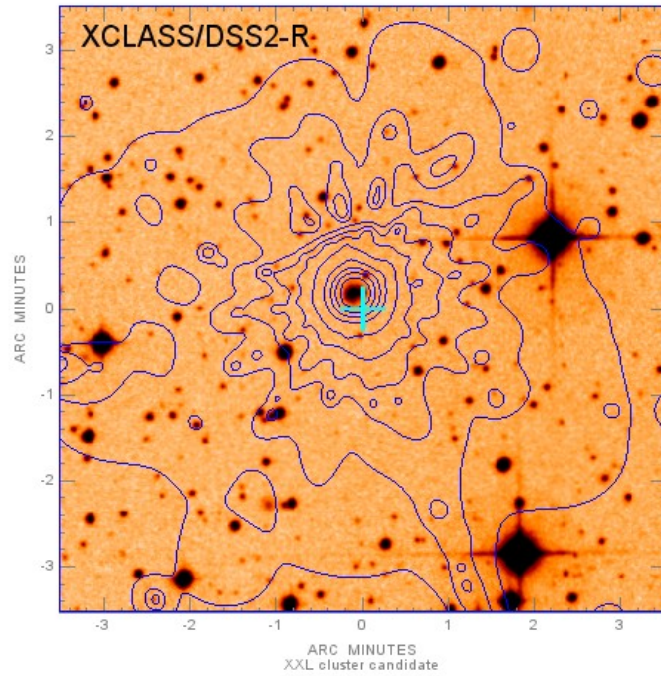
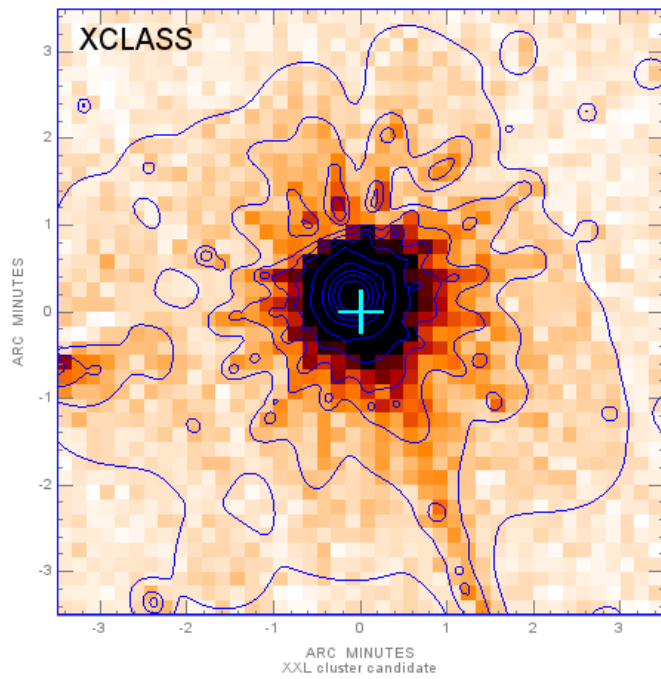
Table 3. The architecture of our custom network which achieved the best performance. Each of the convolutional and dense layers is followed by a ReLU non-linearity with the exception of the final output dense layer which has the softmax for classification.

Layer	Layer type	filter shape / stride	input shape
1	conv	$3 \times 3 \times 64 / (1, 1)$	$356 \times 356 \times 3$
2	max pool	$2 \times 2 / (2, 2)$	$356 \times 356 \times 64$
3	conv	$3 \times 3 \times 32 / (1, 1)$	$178 \times 178 \times 64$
4	max pool	$2 \times 2 / (2, 2)$	$178 \times 178 \times 32$
5	conv	$3 \times 3 \times 32 / (1, 1)$	$89 \times 89 \times 32$
6	max pool	$2 \times 2 / (2, 2)$	$89 \times 89 \times 32$
7	conv	$3 \times 3 \times 32 / (1, 1)$	$45 \times 45 \times 32$
8	max pool	$2 \times 2 / (2, 2)$	$45 \times 45 \times 32$
9	conv	$3 \times 3 \times 32 / (1, 1)$	$23 \times 23 \times 32$
10	max pool	$2 \times 2 / (2, 2)$	$23 \times 23 \times 32$
11	conv	$3 \times 3 \times 32 / (1, 1)$	$12 \times 12 \times 32$
12	max pool	$2 \times 2 / (2, 2)$	$12 \times 12 \times 32$
13	flatten	-	$6 \times 6 \times 32$
14	dense	256	1152
15	dense	2	256

Table 4. Hyperparameters of our custom network and the MobileNet network. The number of iterations, batches yielded during training, is shown for training on the data set classified by experts.

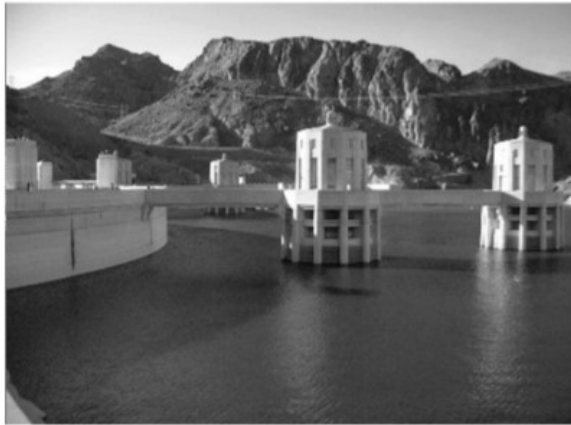
Hyperparameters	Custom net	MobileNet
Batch size	10	20
Iterations	153 000	3 825
Optimizer	SGD	Adadelta
Nest. Momentum	0.90	-
Rho	-	0.95
Initial lr.	0.0001	1.0
lr. decay	10^{-6}	0.95
Minimal lr.	10^{-4}	0.01
lr. red. patience	14	4
lr. red. factor	0.75	0.85
Dense dropout	0.65	0.65
Output activation	softmax	softmax
Loss function	cat. crossentropy	cat. crossentropy
Input image size	356×356	224×224



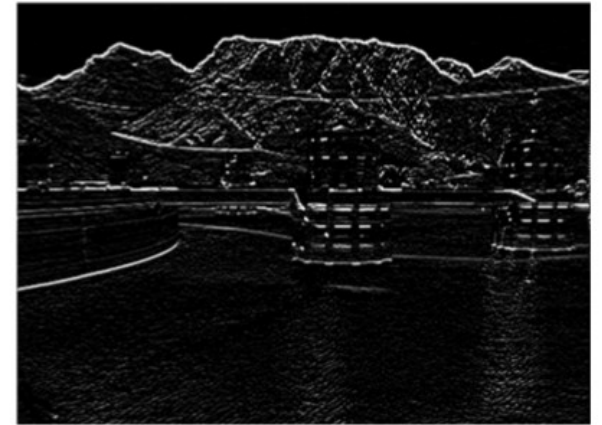


Convolutional filters

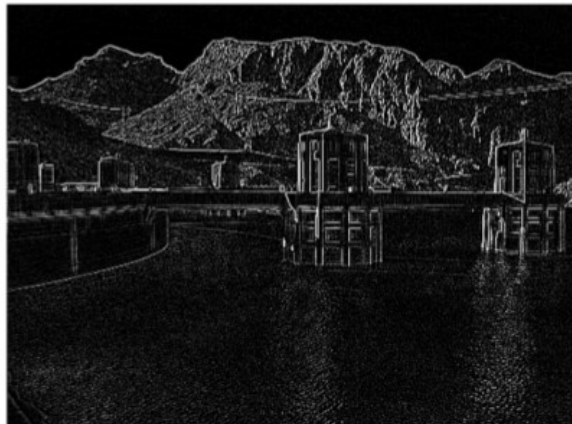
0	0	0
0	1	0
0	0	0



1	2	1
0	0	0
-1	-2	-1



-1	-1	-1
-1	8	-1
-1	-1	-1



-1	0	1
-2	0	2
-1	0	1



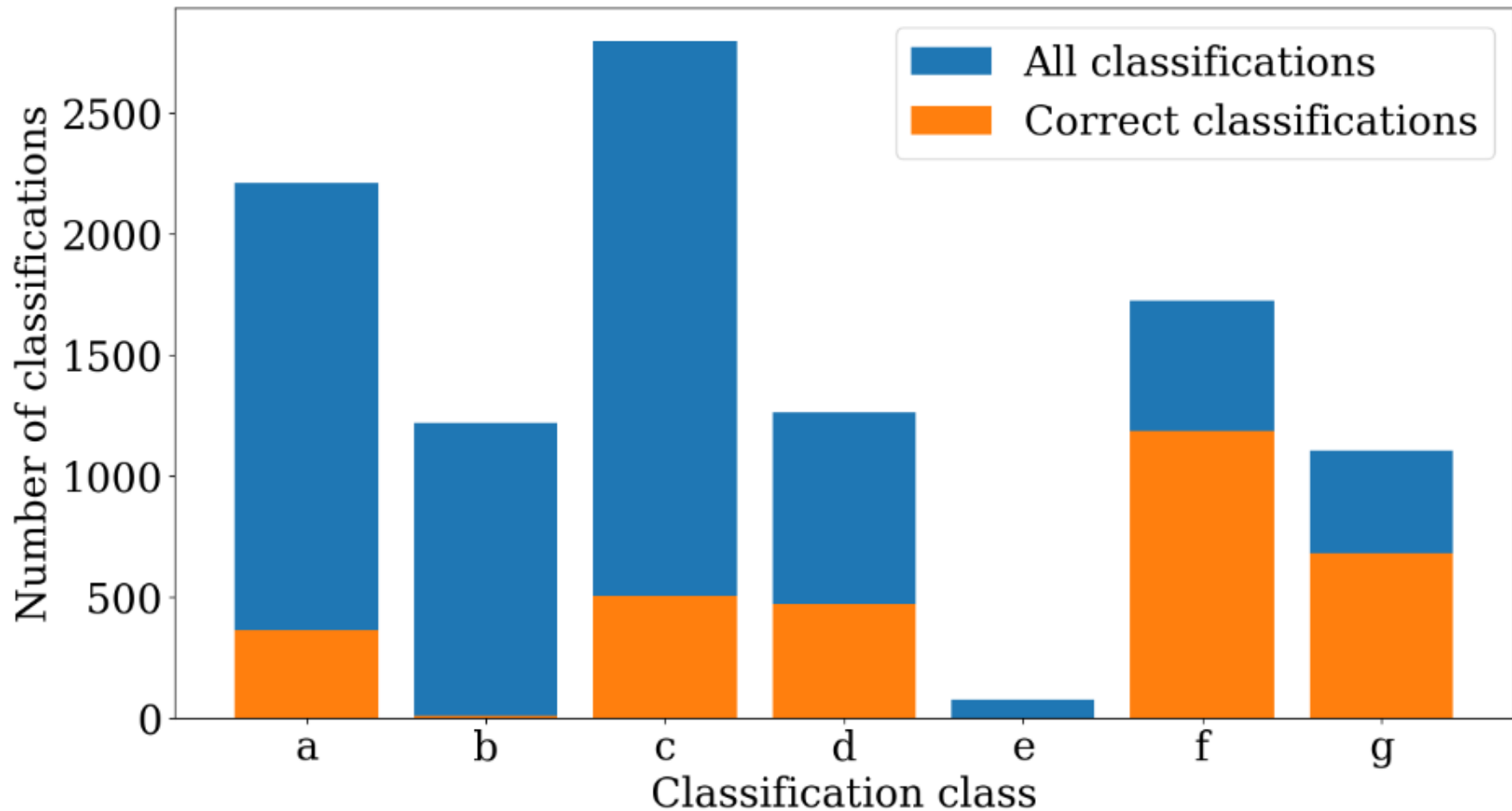


Figure 4.1: Histogram comparing classifications in agreement with experts to all classifications of the Zooniverse volunteers. Coding of answers is following: a) no emission, b) edge, c) point, d) nearby galaxy, e) no optical image, f) nearby galaxy cluster, g) distant galaxy cluster.