Can deep learning help us create better Point Spread Functions, faster?

NSERC-CREATE New Technologies for Canadian Observatories (NTCO) Training Program Research Project

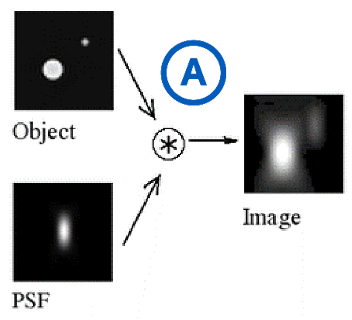
Presented at Canadian Undergraduate Physics Conference, October 2022

NTCO Student: Ashley Ferreira
(University of Waterloo)
Contact: a4ferreira@uwaterloo.ca

NTCO Supervisor: Dr. JJ Kavelaars
(University of Victoria, NRC Herzberg)

Research Project Supervisor: Dr. Wesley Fraser
(University of Victoria, NRC Herzberg)

Introduction



What is a Point Spread Function (PSF)?

- PSFs mathematically describe how point source objects are distorted in an image.
- Images are a convolution between the true object and the PSF as shown in the figure above.

Why are PSFs important to astronomy?

- PSFs are necessary to study any object close to the resolution limit of a telescope with high precision.

What do we need in order to create PSFs?

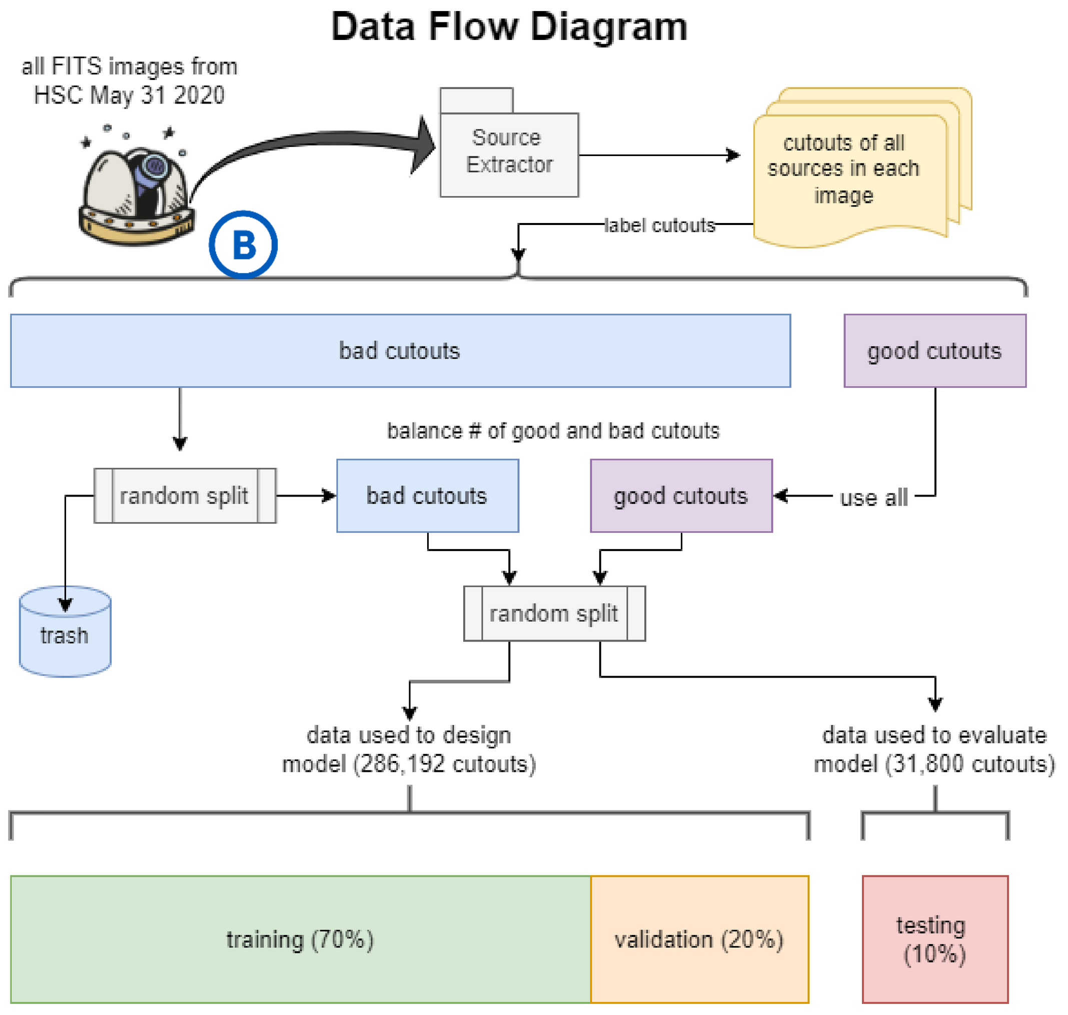
- Examples of point-like sources in the image of interest are needed as inputs to PSF generation software.
- In astronomy, good point-like sources would be stars that are bright, round, and well isolated from other sources.
- The task of selecting these good sources for PSF generation is what this deep learning model has been trained to do.

Data & Methods

For this project, images from 2020 taken by the Hyper Suprime-Cam (HSC) on the Subaru telescope were used as follows:

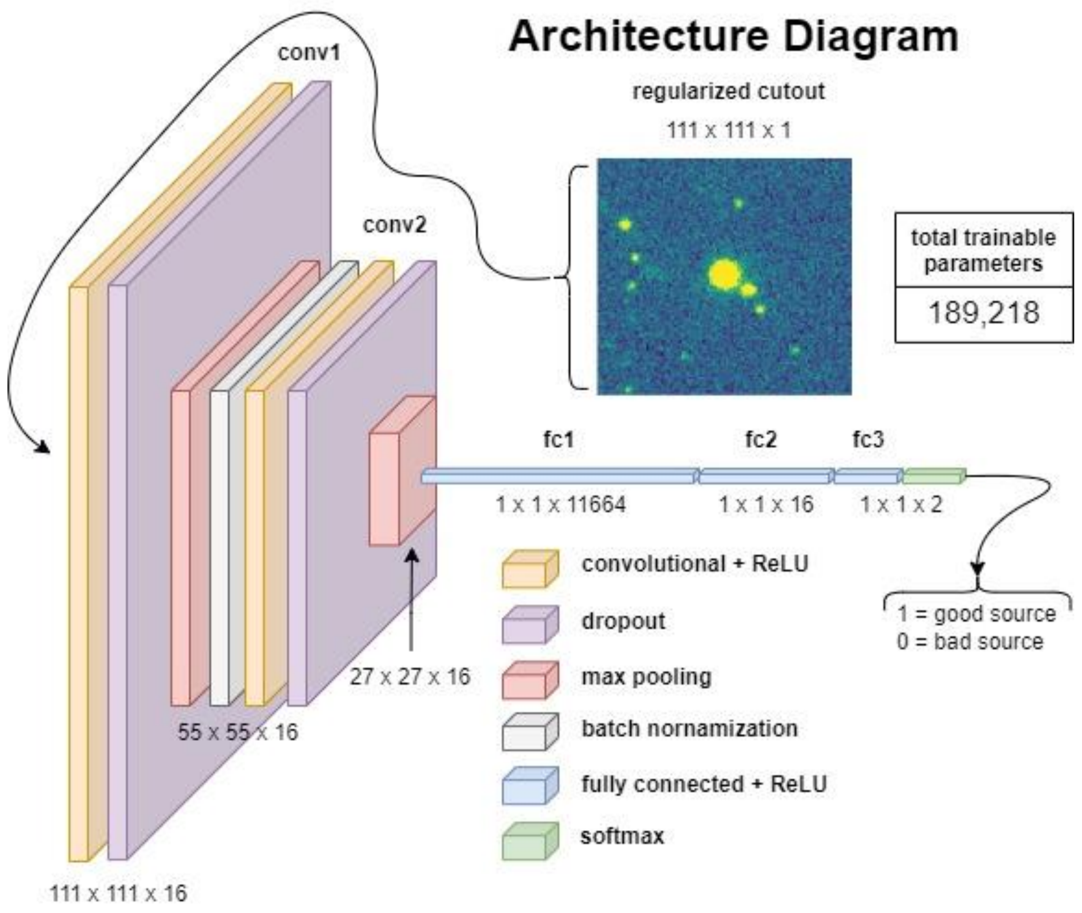
- May 31: to design and evaluate the model.
- May 26: to gauge actual performance in good source selection and the resulting PSF generation.
- For each image, the top 25 sources were selected as those with the lowest flux outside the central source, as inferred by the flux of the most discrepant pixel in the source-PSF residual, and the standard deviation of all residual pixels.
- Of these top 25, the ones which fell in the accepted range of pixel brightness values were deemed good and labelled 1. All other sources were considered bad and labelled 0.
- Using this approach, there are far more bad sources than good ones and so a random selection of bad sources is made such that the 0 and 1 class sizes are equal.

Data Flow Diagram



A simple 2D Convolutional Neural Network (CNN) was developed for this binary classification problem and its structure is shown below.

Convolutional Neural Network Architecture Diagram



Evaluation Metrics

- We are most interested in having a high True Positive rate, a low False Positive rate, and therefore a high good source classification accuracy.
- This is reflected in a precision value close to 100%.

Confusion Matrix Explained

	Predicted 0	Predicted 1
Actual 0	True Negative (TN)	False Positive (FP)
Actual 1	False Negative (FN)	True Positive (TP)

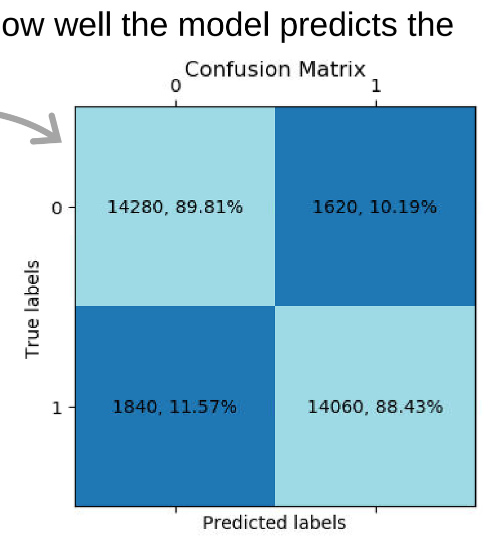
$$\text{Accuracy} = \frac{TP+TN}{(TP+FN+TN+FP)}$$

$$\text{Precision} = \frac{TP}{(TP+FP)} \qquad \text{Recall} = \frac{TP}{(TP+FN)} \qquad \text{Specificity} = \frac{TN}{(TN+FP)}$$

Results

Using the test set we can determine how well the model predicts the labels of an unseen dataset:

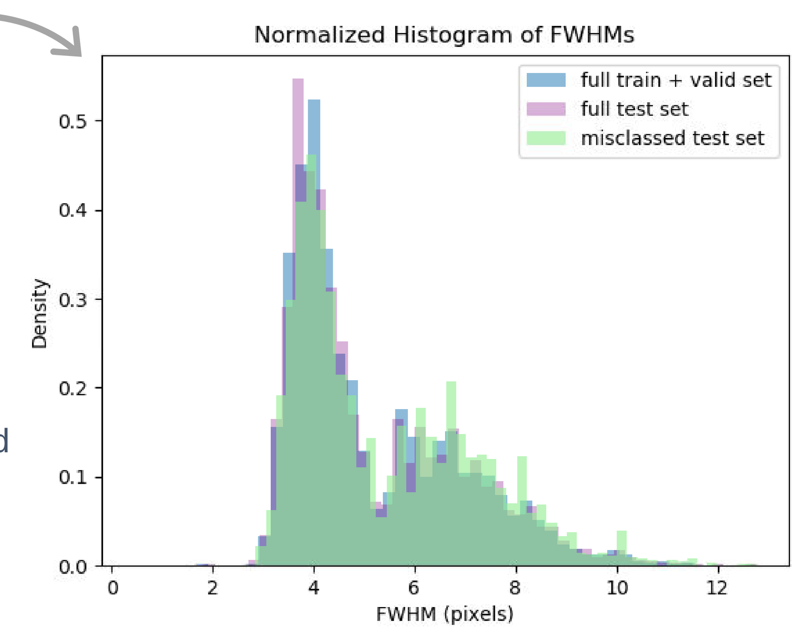
Confusion Matrix:
The accuracy was found to be 89.12% overall.



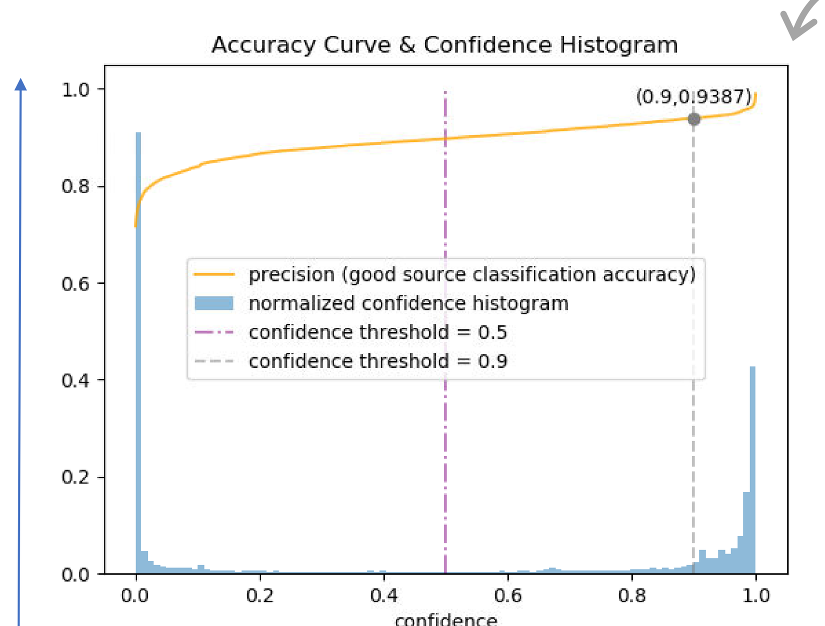
The cutout size 111x111 pixels is used so that images with a variety of Full Width at Half Maximum (FWHM) values can be used with the model.

The differing FWHM values here are due to different seeing conditions but FWHM values are also related to the telescope itself.

FWHM Histogram:
There is no significant, systematic difference in misclassification of sources based on their FWHM.

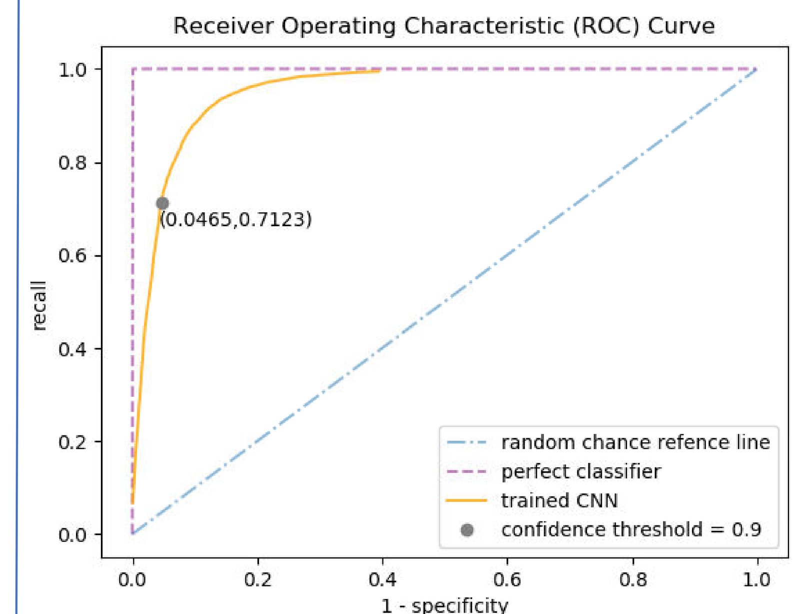


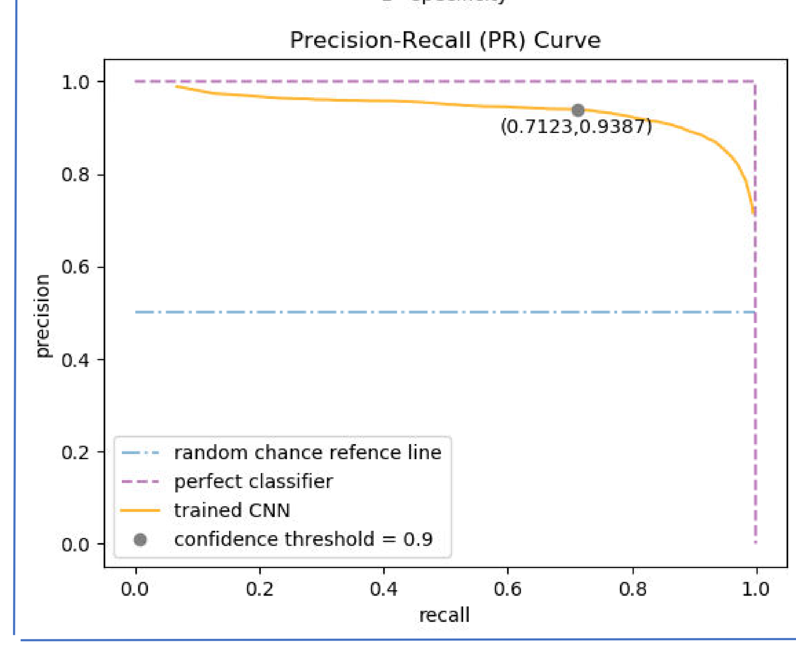
We can raise the confidence threshold beyond which the model labels a source as good. A threshold of 90% was adopted such that we can achieve a precision of 93.87% while still having a significant number of sources classified as good.



Accuracy Curve:
As the good source confidence threshold is increased, we see increased precision.

Confidence Histogram:
The model is decisive.



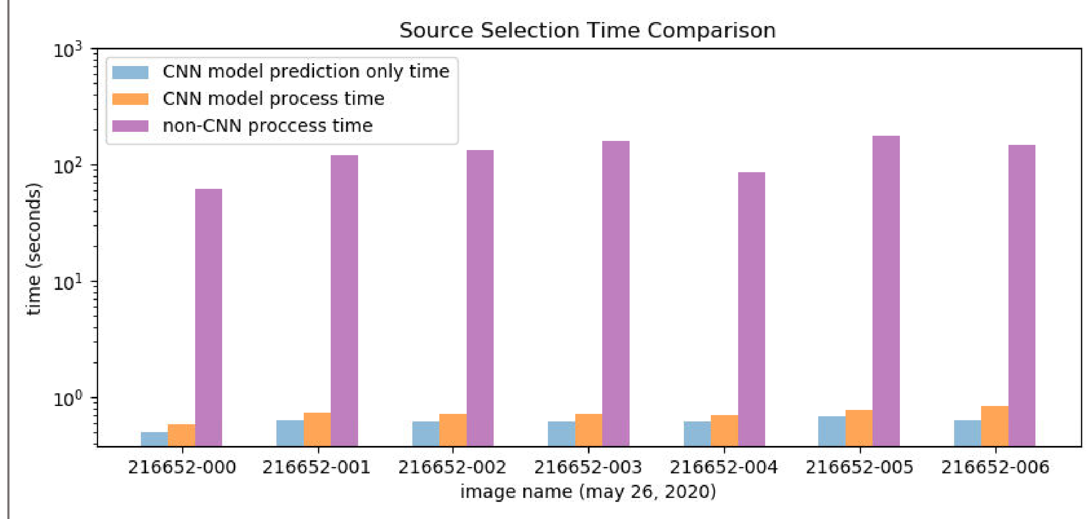


ROC Curve + PR Curve:
The behavior is as we would hope for, not perfect but far better than random guessing.

Increasing good source confidence threshold

Conclusion

- Once the model is trained, the CNN method takes only ~6% the CPU time of the non-CNN method.



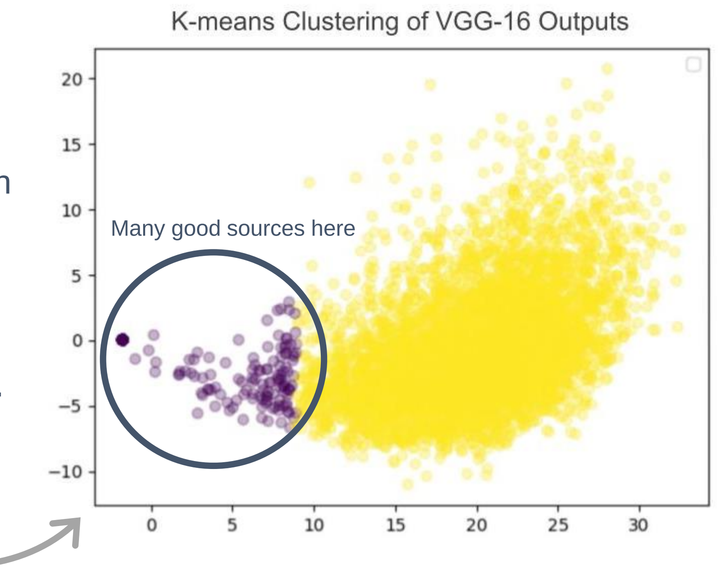
Next Steps

- **Make model public:** model to be incorporated into PSF generation software TRIPPy as option to help users with source selection. A platform such as OpenML may also be used to share data, model, and results with others interested in similar work.

Possible Improvements

- **Model + hyperparameters:** tracking experiments with a tool such as Weights and Biases in order to find an optimal model could improve results.
- **Improving labelled data:** data labels used are not perfect and so a combination of manual relabeling and using the assistance of clustering will improve results. Comparing k-means clustering results or outputs from a One-Class Support Vector Machine to labelled results could highlight outliers in need of relabeling or even provide an alternate solution to entirely.

Clustering:
Outputs of pretrained CNN image classifier with final layer removed are used for k-means clustering into two classes (purple and yellow).



Future Directions

- **Apply to other telescopes:** this model should be able to work with data from other telescopes, however methods such as down/up sampling might need to be used it may will likely be needed to allow for 111x111 inputs.
- **Increasing scope of model:** using similar data it may be possible to have a model find the sources in an image directly and/or generate a PSF as the output instead of using pre-made cutouts of sources and just outputting a prediction on if they should be used for PSF generation or not. Methods are also currently being explored to perform PSF deconvolution on images directly using CNNs.

GitHub **Repository:** CNN model as well as code used to develop and test it

Google Slides **Examples:** Z-scaled examples of good source selection and resulting PSF lookup tables

References & Acknowledgements

A huge thanks to my supervisors, especially Dr. Fraser for all his support on this project as well as NTCO for the opportunity to be involved in this work.

(A) Source: Convolution Illustrated eng on Wikimedia Commons

(B) Source: Space Observatory Icon from Good Stuff No Nonesense

SExtractor to extract sources from FITS images
TRIPPy to generate PSFs
CANFAR used as compute environment

CANFAR

NTCO UNIVERSITY OF WATERLOO