**The Current State and Future Prospect of Machine Learning in Astronomy**

Alexander Wei, Tony Xia, Ryan Zhang

SES4UI

Michael Burns

June 14, 2019

**Table of Contents**

1. **Abstract**
2. **Introduction**
3. **Applications of Neural Networks**
   1. **CNN**
   2. **GAN**
4. **Our Implementation**
5. **Our Results**
6. **Possible Improvements**
7. **Future Prospect**
8. **Conclusion**
9. **Abstract**

In the following paper, we describe the needs and applications of machine learning in astronomy, both today and in the future. Artificial intelligence has arisen to become an essential technique in the study of astronomy, and its applications will be dissected and reviewed.. Thereafter, we constructed a model based on a convolutional neural network to classify various constellations. Our model resulted in a relatively accurate prediction value of 80 percent, which can only be improved with more accurate and diverse data. In the end, we discussed future applications of machine learning in astronomy, including image denoising, pattern recognition, and even the creation of mathematical models to predict physical phenomena.

1. **Introduction**

In 1958, Frank Rosenblatt created the perceptron, a binary classifier that, in essence, output yes or no in response to an input. This machine was modeled after the human neuron, the very thing that creates our consciousness and the thoughts that define humanity. The simple concept of creating a machine that learns like humans is one of the largest fields of today, simply because there is so much data to process, too much for humans to comprehend. One field that has an influx of data without a meaningful interpretation is the field of astronomy. With advancing imaging technology, images of the universe is taken at an ever-increasing rate. Large Synoptic Survey Telescope, a new telescope that is scheduled to start working in 2020, is estimated to produce 500 petabytes (51017 bytes) of data over 10 years. It is impossible to process and analyze all the data by humans, and thus it is very important to automate the process. Due to the intrinsic variability of the images, artificial intelligence is the only and the most natural way to approach this problem. In the field of astronomy, many studies have already achieved surprising results by applying neural networks. For example, in 2017, a technique developed by Stanford allows artificial intelligence to detect and analyze gravitational lensing in images. It is estimated that this technique is 10 million times faster than the traditional method which requires up to a month to analyze one image. Artificial intelligence is bringing astronomy into a new era, where computers augment the ways that researchers can consume and understand data.

This report will tackle the topic of neural networks and its applications in the data-driven field. To prove the applicability of neural networks, we have endeavoured to create a model that classifies and predicts constellations based on an image. Our model, being a proof-of-concept, only classifies a limited number of constellations. This limit is easily exceeded with more image data, however.

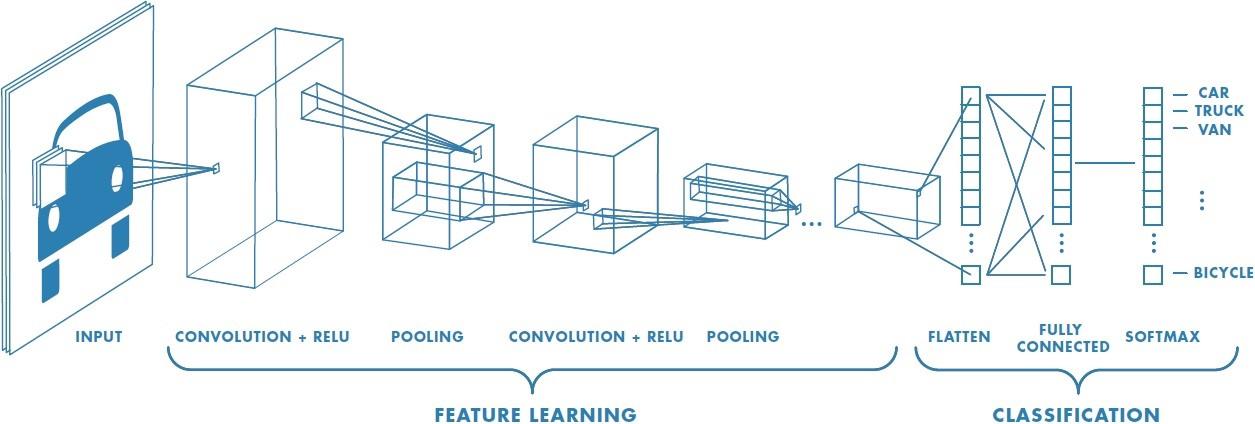
Object classification is not the only thing machine learning can offer to astronomy. Besides the implementation of a neural network, in this report, we also explore different types of neural networks that are applicable in machine learning, namely the convolutional neural network and the generative adversarial network, to unveil the power artificial intelligence has brought about to the area of astronomy. Furthermore, we discussed some future possibilities neural networks can bring and the potential impact neural networks could have on not only astronomy, but all science areas as a whole by presenting a new study done by a group from Switzerland.

1. **Applications of Neural Networks**

### To understand the applications of NNs, first we must understand the theory behind neural networks. A neural network is a series of neurons, also known as nodes, that take in input and output a value. A neuron can be thought of as a linear equation: y = mx + b, where x is the input, m is the weight assigned to the input, b is the bias of the node, and y is the output of the node. Normally, the inputs and outputs are normalized to values between 0 and 1, to remove any natural biases for differing magnitudes of data. A prime example of this normalization is the division of the pixel brightness values by 255, which is the max brightness value, putting the resulting value between 0 and 1. A series of these neurons are strung together to form a network. The network is trained with respect to a loss function, a function that calculates the error between the predicted value and the actual value. Imagine a model that predicts the housing prices of a city like Toronto. It would take an input of years, and output a predicted price for an average apartment. For example, if the price of the apartment after 4 years is 5000, and the model predicts a price of 4000, the difference in the values would be 1000, and that is what the loss function would output. In more complicated situations, the loss function would be different, but would still output the “wrongness” of the model. This value is then used to tweak the weights and biases to train the model. These processes together create a network of nodes where the weights and biases are tuned to give predictions. Predictions can consist of discriminating between various classes, known as a classification problem, predicting a quantity based on an input, known as a regression problem, or even denoising data, useful for data mining.

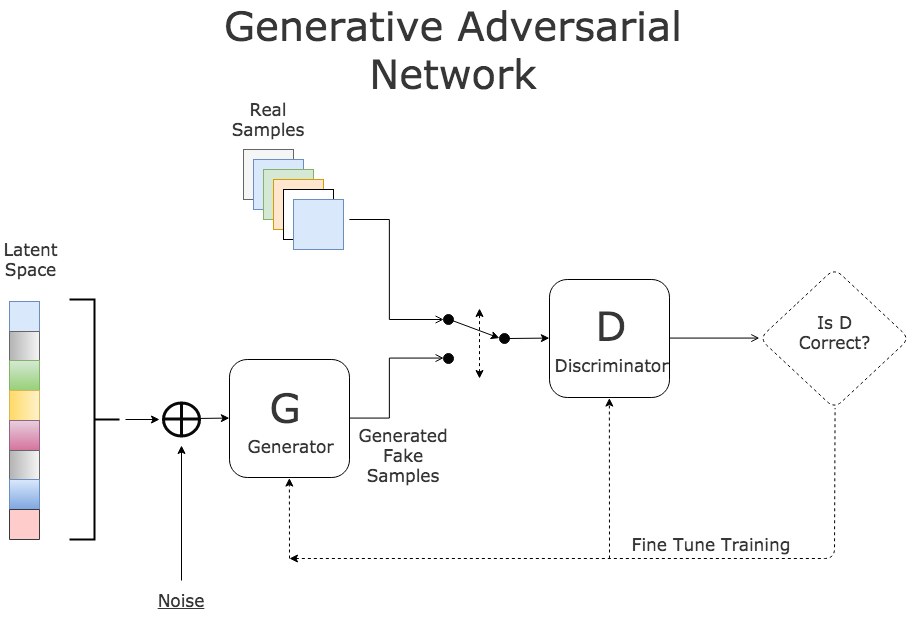
**a. CNN**

### The CNN, or Convolutional Neural Network, is a derivative of the Artificial Neural Network. This network is different due to convolution, a process that groups pixels together so that the network can formulate more abstract predictions and learn not only the pixel values, but also the features that are important in classifying the image. A CNN is used primarily in image classification, as it is translationally independent. In astronomy, the CNN can be used to classify different images, even out classing human prediction. The uses for this application are nearly limitless, from identifying objects within images, to finding new and interesting points of interest within images. This technique can help alleviate the human load of scanning satellite images for new, undiscovered objects.

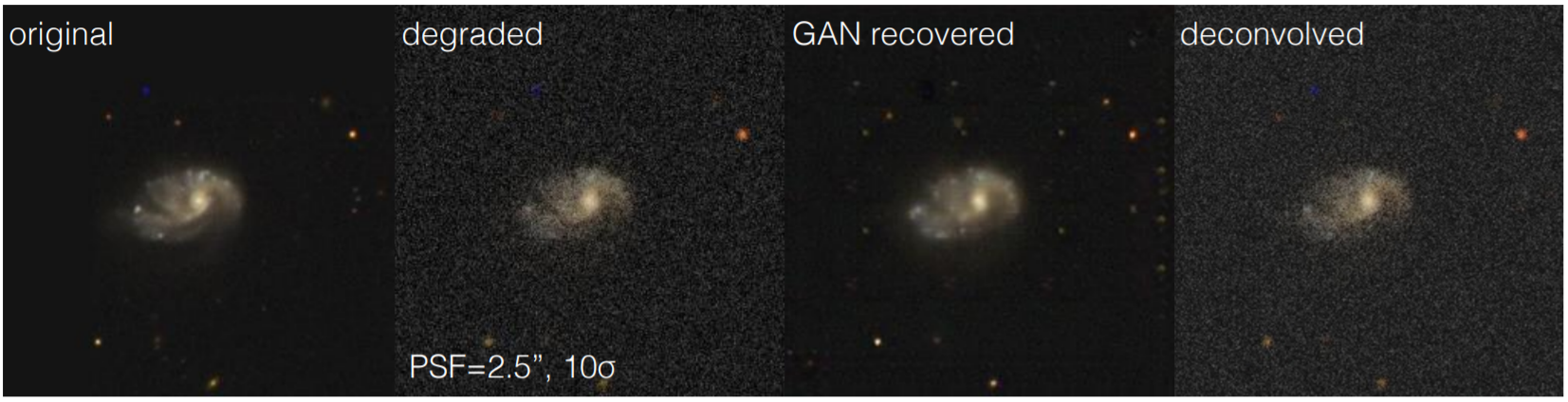


**b. GAN**

The generative adversarial network (GAN), is another type of neural network that could aid astronomers. Rather than one single neural network, a generative adversarial network consists of two individual neural networks, one responsible for generating images, one responsible for analyzing images. The generative network generates new images, and the discriminative network, after being trained with real images, evaluates and determines if an image is generated by the generative network. While the generative network aims to increase the discriminative network’s error rate, the discriminative network aims to reduce its error rate. This competition between two networks gradually improves the generative neural network’s ability to create new images, which in turn is a really effective way of generating photo-realistic images with computers.



The generative neural network can be applied to repair and enhance images with low signal-to-noise ratio. With generative adversarial network, astronomers are able to recover images that are not possible with traditional methods. The following diagram is a comparison between GAN and conventional deconvolution method.[[1]](#footnote-0) The original image is degraded before fed into the GAN network and the deconvolution algorithm, and as it can be seen, the GAN recovered image is a lot clearer than the deconvolved image.



As more and more images come in from various telescopes around the world, this technique can be applied to a wide range of current and future data. With this technique, we will be able to study astrophysical objects more easily.

1. **Our Implementation**

### Our implementation utilizes the comprehensive Tensorflow library, which encompasses Keras, a Python-based deep learning library. The process consists of getting images, which came from Google. Since constellations are captured in various rotations, we took our images and saved an additional 359 rotated versions of the image, one for each degree of the circle. We then converted the images into grayscale versions and also shrunk the image down to a 256 by 256 pixel png image for the NN to process. We then take each pixel brightness value, represented as integers from 0 to 255, and divide by 255. This normalizes the range of numbers into numbers between 0 and 1, the required format needed for the NN input.

Using the pathlib library for Python, we can easily extract all the image files in our directory in order to use for training. Our current model was trained with 35 epochs with 30 steps per epoch. Every step per epoch randomly selects 32 images from our dataset and trains with it. Thus, every epoch utilizes 960 random images from our collection.

Additionally, we used max pooling to shrink the size of our array, making the values more abstract and also reducing the number of trainable parameters. This “abstraction” generalizes the image array values more in hopes to prevent overfitting. Our current model contains 68 272 389 trainable parameters. A potential reason behind this unusually high number is because our images are saved with dimensions 256 by 256, much larger than what might be used in different neural network datasets.

Overfitting is an issue when the model fits too closely for the amount of data given. As mentioned earlier, we used max pooling to reduce the number of parameters needed to be fit. In the particular case of our study, a prototype of our model was “memorizing the locations” of the stars rather than searching for defining patterns that make up a constellation. As a result, our older model was not actually “learning” anything so when our new test data was introduced to the model, it was unable to successfully classify them. Reducing the number of trainable parameters by max pooling was crucial for our model to function effectively and fix this issue.

1. **Our Results**

Our current model can be further improved with more raw data for training. Because our images were pulled merely from Google searches, we lacked quality sets of each of the constellations. Allowing the model to train with greater amounts of data will vastly improve accuracy of predictions. Our model trained with 8 of each constellation, with the exception of Leo, which we had 4 images of. Rather than gathering massive amounts of images by hand, we used all of our rotated images for training. This meant that we had a rather weak dataset because of the small number of unique images. Obviously, implementing the training set like such is nonoptimal and we do not recommend doing so.

One particular method of obtaining more images for the training of the model was to translate the images. By positioning the constellation in both different angles and in different parts of the image, we can introduce more variance in the training set, potentially yielding better results.

Currently, our model hovers around a 95% accuracy during the training sequence and is able to predict constellations from our test set moderately well. After training the model with about 35 epochs and 30 steps per epoch, the model can successfully classify approximately 4 out of 5 constellations. To improve the accuracy, we recommend obtaining more images and also implementing some horizontal and vertical translations into the training dataset.

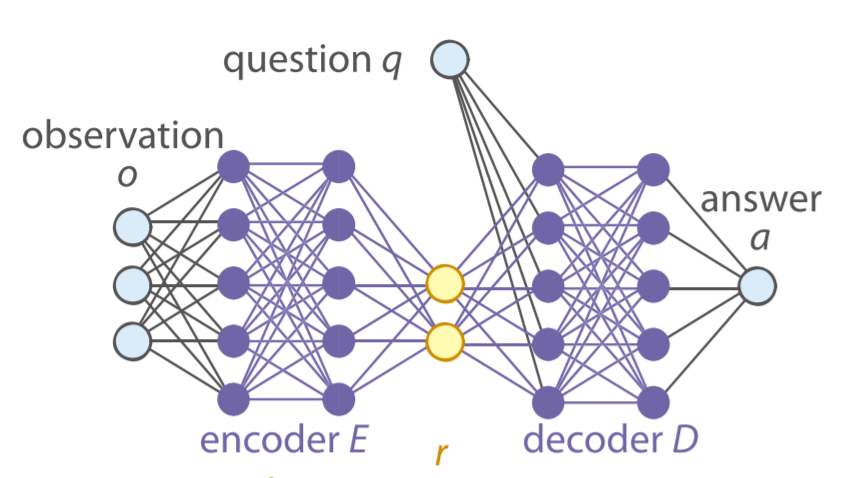
Our model is able to predict some constellations better than others. Currently, we only have implemented Orion, Ursa Major, Cassiopeia, Gemini, and Leo, although it is not more difficult to add more constellations. Constellations such as Orion and Gemini are easier to predict because they contain more defining stars in the constellation. Our model tends to struggle with identifying smaller constellations such as Cassiopeia and Leo simply because they have fewer stars compared to others.

1. **Possible Improvements**

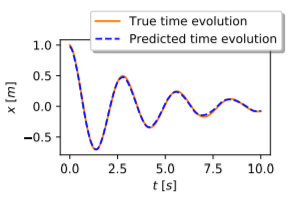
Through testing and use of the model, there are definite flaws that can be fixed in the future. These flaws consist of data variance, data noise, and model design. Data variance occurs when the data is the same or very similar, which occurs with our dataset, since there are essentially 360 copies of each image. To fix this problem, we need to introduce more images, as well as different perspectives of constellations, whether through a deep-space telescope, or through an amateur photo. Data noise is a problem in training due to random variables that take up precious training time, and can also introduce false positives that the model can remember. This would be solved through proper feature selection, where the important portions of each image are fed into the model. This method was attempted by us by turning the images into pure black and white images, with no gradients. While this did reduce the file sizes and network load, it cut out many stars that were integral to the proper assessment of the constellation. To combat this in the future, one would need to create a dynamic program that would adjust for overall brightness of each image, and cut noise. Our model design could also be improved using techniques like those employed in ResNet, which includes novel architectural changes called “skip connections,” where data from one layer is fed not just to the layer after, but to all subsequent layers. This would allow for better analysis of what is considered to be a sparsely populated image, and result in much greater accuracy.

1. **Future Prospects**

A group of researchers from the Institute for Theoretical Physics, ETH Zurich developed a neural network that can discover physical laws with only experimental data. This neural network consists of two sections: an encoder and a decoder. Each of the sections is a feed-forward neural network. The two sections simulate how a scientist would develop physical laws. First, the input layer gathers data and feeds the data into multiple layers of neurons. Each of these neurons processes the data, and feeds the output into the next layer. After multiple layers, the data is compressed down to a minimal representation of variables (such a neural network is also called a “feed-forward neural network”). This minimal representation, or latent representation, is what the neural network derives as the “physical law”.

The second section of the neural network is the encoder. Similar to the encoder, the decoder is also a feed-forward network. It interprets the latent representation and applies it to solve a specific question. The overall architecture of the neural network looks like the following:

Testing with this neural network, the networks was able to do the following:



1. discover a correct model for the motion of a damped pendulum,
2. derive and exploit the conservation laws,
3. determine the correct degree of freedom of a quantum state,
4. discover the heliocentric model of the solar system.

The success of this model raises a problem: can machines help humans simplify the existing theories? Are the theories we currently have the most natural approach to explaining natural phenomena? Will artificial intelligence help humans with the development of scientific theories by presenting the optimal model for a study subject, thereby eliminating the need for humans to devise physical laws?

With our current understanding of neural network and deep learning, we are not be able to answer these questions, but it is potentially where we are going to go.

1. **Conclusion**

In conclusion, being able to apply machine learning techniques in astronomy studies has become an essential skill and is increasing in importance. Artificial intelligence can now help astronomers study objects in the universe more efficiently and effectively than ever before. As shown by our example program, a neural network can achieve a reasonable accuracy at classifying images even without a lot of data. With a larger training dataset, a much higher accuracy can be acquired, thereby making this model more viable. Besides classifying images, neural networks can also help astronomers repair images, remove image noise, and make studying of deep-sky objects easier. In the future, artificial intelligence could potentially change how people do scientific research, as it provides people with an easy way to model physical phenomena.

References

(n.d.). Retrieved from https://blogs.umass.edu/brain-wars/1957-the-birth-of-cognitive-science/the-perceptron-a-perceiving-and-recognizing-automaton/

Das, S., & Das, S. (2017, November 16). CNN Architectures: LeNet, AlexNet, VGG, GoogLeNet, ResNet and more .... Retrieved from https://medium.com/@sidereal/cnns-architectures-lenet-alexnet-vgg-googlenet-resnet-and-more-666091488df5

Hezaveh, Y. D., Levasseur, L. P., & Marshall, P. J. (2017, 08). Fast automated analysis of strong gravitational lenses with convolutional neural networks. *Nature,* *548*(7669), 555-557. doi:10.1038/nature23463

Iten, R., Metger, T., Wilming, H., Del Rio, L., & Renner, R. (2018). Discovering physical concepts with neural networks. arXiv preprint arXiv:1807.10300.

Ivezić, Ž, Kahn, S. M., Tyson, J. A., Abel, B., Acosta, E., Allsman, R., . . . Zhan, H. (2019). LSST: From Science Drivers to Reference Design and Anticipated Data Products. The Astrophysical Journal,873(2), 111. doi:10.3847/1538-4357/ab042c

Kenton, W. (2019, April 15). Overfitting. Retrieved from https://www.investopedia.com/terms/o/overfitting.asp

Lamb, E. (n.d.). Studying the stars with machine learning. Retrieved from https://www.symmetrymagazine.org/article/studying-the-stars-with-machine-learning

Max-pooling / Pooling. (n.d.). Retrieved from https://computersciencewiki.org/index.php/Max-pooling\_/\_Pooling

Schawinski, K., Zhang, C., Zhang, H., Fowler, L., & Santhanam, G. K. (2017, 01). Generative Adversarial Networks recover features in astrophysical images of galaxies beyond the deconvolution limit. *Monthly Notices of the Royal Astronomical Society: Letters*. doi:10.1093/mnrasl/slx008

1. The diagram is the result of other people’s work. We do not own the images in the diagram. They are placed here just as an example. (Schawinski) [↑](#footnote-ref-0)