Journal of Research High School

Vol. 2024 (12) 560 – 570
ISSN 2688-3651 [online]

Performing Spectroscopic Parallax via Machine Learning Classification of Stars in the Morgan-Keenan System to Efficiently Calculate the Distance to Stars

Jaymin Ding[1] *

[1]Rye Country Day School, Rye, NY, USA
*Corresponding Author: jtding43@gmail.com

Advisor: Dr. Mary Krasovec, mary_krasovec@ryecountryday.org

Received May 31, 2024; Revised September 27, 2024; Accepted October 11, 2024

Abstract

The distance to a star is one of the most fundamental pieces of information that can be calculated about it. Currently, several methods exist for calculating distance; however, many of them are quite inefficient and prone to error. To combat this, one could classify to find its properties and subsequently use the classification to calculate the distance. Tools exist to classify stars, but lack in terms of accuracy and ease of use. To make measuring distance more efficient and accessible for professional and amateur astronomers alike, we developed a novel method for measuring stellar distances using spectroscopic parallax and machine learning. The procedure of this project goes in two steps: first, to utilize Google's MobileNetV2 to identify a star's spectral type on the Morgan-Keenan System with 97.9% accuracy, and second, to use these classifications to calculate the distance to the star with relatively low error. Classifying a star into a Morgan-Keenan System provides data on the temperature and color of a star through spectral class and size, luminosity, and evolutionary phase through the luminosity class, and by combining this with Hertzsprung-Russell Diagram data, we may calculate the absolute magnitude, or how bright the star is at a standardized distance from Earth. Using the image itself, the apparent magnitude, or how bright the star looks from Earth, can be calculated. The distance can then be calculated by comparing the absolute and apparent magnitude using a method called spectroscopic parallax. Applying computations to this method makes large telescopes much more cost-efficient.

Keywords: Spectroscopic parallax, Morgan-Keenan System, Hertzsprung-Russell Diagram, Star Distance, Machine Learning, Transfer Learning

1. Introduction

1.1 Background

Measuring the distance to a star is important in understanding the expansion of the universe, mapping the structure of the Milky Way, and planning future space missions. Current methods of measuring distance usually involve long timespans or a range of uncertainties (Carroll and Ostlie, 2018), and as such, large inefficiencies are created in the process. To combat this, the goal of this project was to create an efficient and accurate way to measure distance using computational analysis of images. The most effective way to achieve this was through using computer vision to visually analyze an image of a star and compare its intrinsic brightness to its apparent brightness to determine distance.

1.2 The Morgan-Keenan System

The Morgan-Keenan System is a method currently used by astronomers to classify stars. There are three main aspects of the classification in the Morgan-Keenan System: the spectral class, spectral subclass, and the luminosity class. The spectral class of a star is divided into categories that consist of O, B, A, F, G, K, and M, and from it, one

Vol. 2024 (12) 560 – 570
ISSN 2688-3651 [online]

can get the temperature and color of a star. As a star goes from O to M, it goes from being bluer and hotter to being redder and cooler. The spectral subclass further divides a spectral class from categories 0 to 9. The properties of a star that falls into a spectral class are shown in Figure .The luminosity class of a star is mainly divided into categories I, II, III, IV, and V, with I being further divided into Ia or Ib. Luminosity classes can provide information on a star's size, luminosity, and evolutionary phase. A general description of a star that falls into each luminosity class is shown in Table 1.

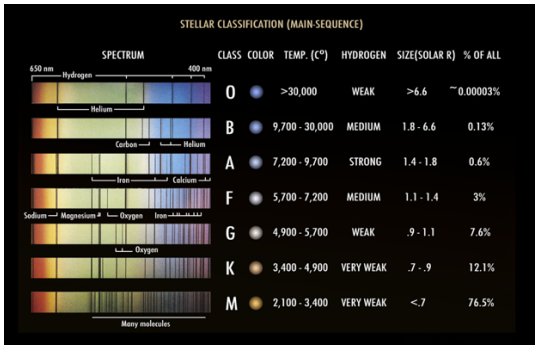


Figure 1. A chart of the Morgan-Keenan System spectral classes and the range of properties for stars falling into specific categories (Budassi, 2024).

Table 1. A chart of the Morgan-Keenan System luminosity classes and a general description of stars in each type of luminosity class (Brau, 2016).

Class	Description
Ia	Bright Supergiants
Ib	Supergiants
II	Bright giants
III	Giants
IV	Subgiants
V	Main-sequence stars and dwarfs

1.3 The Hertzsprung-Russell Diagram

The Hertzsprung-Russell Diagram plots the luminosity against the temperature in a decreasing direction for a number of stars. An example is shown in Figure 2.

The Hertzsprung-Russell Diagram is highly valuable in the comparison of stars, especially as stars tend to clump together by luminosity class, as can be observed in Figure . An interesting trend to note is that main sequence stars tend to follow a path diagonally from the top left to the bottom right of the diagram. The data on the Hertzsprung-Russell Diagram can be used to approximate the absolute magnitude of a star given both the spectral class and the luminosity class, which work because they are measures of temperature and luminosity, respectively.

1.4 The Stellar (Trigonometric) Parallax Method of Measuring Distance

One method for measuring distance that astronomers currently use is known as stellar parallax. A diagram showing how it works is shown in Figure (Amos, 2016).

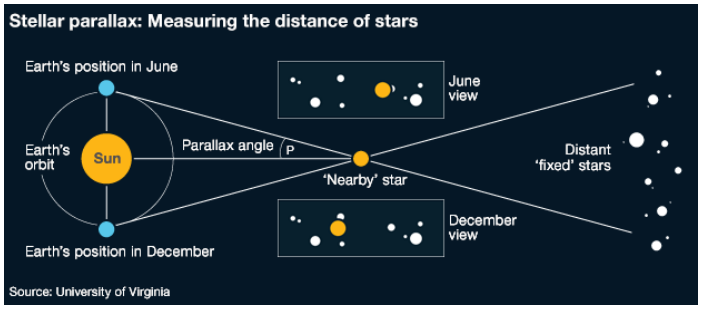


Figure 3. An illustration of the stellar parallax method.

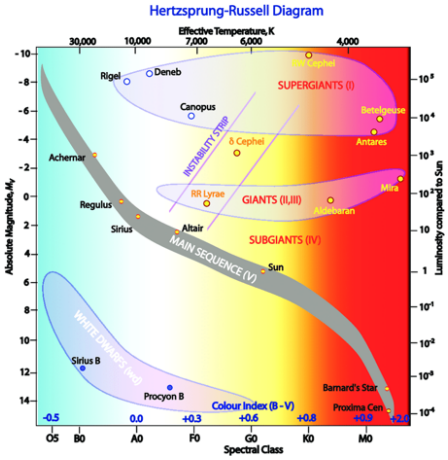


Figure 2. An example of the Hertzsprung-Russell Diagram where color and spectral class are plotted on the *x*-axis, and luminosity and absolute magnitude are plotted on the *y*-axis (Althaus et al., 2010).

When Earth makes half an orbit around the sun, the position of a nearby star shifts ever so slightly. Such a shift can be measured using high-precision equipment from Earth; however, such a small shift means that it becomes difficult to measure distances using stellar parallax beyond 100 parsecs (pc) from Earth. In addition, the method takes six months to complete one measurement.

Vol. 2024 (12) 560 – 570
ISSN 2688-3651 [online]

The distance to the star in pc, denoted by d, can be calculated from the parallax angle in arcseconds (as), denoted by p'', using Equation (1) (Carroll & Ostlie, 2018):

$$d = \frac{1}{p''} \quad (1)$$

The parallax angle is also commonly measured in terms of milliarcseconds (mas), in which case the distance to the star in parsecs can be calculated using Equation (2):

$$d = \frac{1000}{p''} \quad (2)$$

1.5 Literature Review

Existing work has shown to be able to classify stars using different types of machine learning models; however, many come with some caveats. First, these tools are only able to classify stars and cannot use these classifications to then find the distance to them. Second, most of them only reach the 70-80% accuracy range (Dafonte et al., 2020; Hungund, 2020; Kyritsis et al., 2022). Third, many of these tools require the actual stellar spectra as inputs, meaning that individual values such as redshift z have to be plugged in for classification. This can mean that over 20 values have to be inputted for a single classification. In addition, no current work has been found that utilizes machine learning to directly find the distance to stars. As such, the usefulness of a tool that can efficiently and accurately classify stars and find the distance to them is evident.

2. Materials and Methods

2.1 Data Collection

A large dataset is required to create a machine-learning model. The first step in compiling a large dataset of stars was to find the names of many, and so a list of 430 stars was gathered from the List of International Astronomical Union (IAU)-approved Star Names (International Astronomical Union | IAU, 2015). Once the list of star names was gathered, the Set of Identifications, Measurements and Bibliography for Astronomical Data (SIMBAD) online astronomical database was queried for the spectral type and parallax angles in mas for each star. The parallax angles were subsequently converted into distance by finding 1000 multiplied by the multiplicative inverse of the parallax angle. The spectral type was also parsed into a spectral class, spectral subclass, and luminosity class. Then, the HiPS2FiTS cutout API (Hips2fits - Fast Generation of FITS Cutouts from HiPS Datasets, n.d.), which can provide an image of a celestial object given a name, was queried for images of each star in the CDS/P/DSS2/color Hierarchical Progressive Surveys (HiPS) survey, which is similar to a color scheme. This ensured that the images would come out in colors representative of how humans view the stars, so an O-type star would be blue, an M-type star would be red, and a B-type star would be bluer than an A-type star but not more than an O-type star. Selected data entries are shown in Figure 4 and Table 2.

Table 2. An example of the properties of 5 stars from the dataset of 430 stars.

Star Name	Spectral Class	Spectral Subclass	Luminosity Class	Distance (parsecs)
Alnitak	O	9	I	225.73
Betelgeuse	M	1	I	152.67
Maia	B	8	III	117.51
Taiyangshou	K	0	III	60.834
Vega	A	0	V	7.6795

Vol. 2024 (12) 560 – 570
ISSN 2688-3651 [online]

Alnitak *Betelgeuse* *Maia* *Taiyangshou* *Vega*

Figure 4. An example of 5 stars from the dataset of 430 stars in the CDS/P/DSS2/color HiPS survey (Wenger et al., 2000).

For all queries, the Python programming language, along with the Astroquery, Selenium, and Requests libraries (Chandra and Varnasi, 2015; Ginsburg et al., 2019; Price-Whelan et al., 2018), were used to query the databases and perform web scraping in order to successfully gather the images.

2.2 Data Augmentation

It was decided that using image augmentation, or an expansion of the dataset, would be useful in improving the accuracy of the machine learning model. The augmentation served a twofold purpose. First, it generally expanded the image dataset to an adequate amount. The original data came in a very unbalanced shape: whereas there were a few hundred stars for certain spectral classes, other spectral classes, such as O, had only six stars. As such, by performing image augmentation, sufficient amounts of images were made for each category. Second, by rotating and flipping images, it encouraged the machine learning model to focus more on the star at the center of the image, which does not change very much due to it being approximately spherical and symmetrical, instead of the background of the star, which does not matter in its classification. Being able to recognize stars in various orientation was very beneficial in improving the accuracy of the model, which is the main reason why these particular methods of augmentation were chosen as opposed to other methods, such as cropping or changing colors. The geometric operations performed on the images were combinations of rotating the images by varying multiples of 90° and flipping them over the x or y-axis. The Python programming language, along with the OpenCV, Pillow, and Albumentations libraries (Bradski, 2000; Buslaev et al., 2020; Clark, 2015; Kalinin, 2018), were used to read images and perform geometric operations on them to expand the dataset. An example of how images were augmented is shown in Figure 5 below.

Original *Rotated 90° Counterclockwise (CCW)* *Reflected across the x-axis* *Rotated 90° CCW and reflected across the x-axis*

Figure 5. An image of the star Alnitak with various geometric operations performed upon it

2.3 Machine Learning Model Creation

Once the dataset was expanded, three separate machine-learning models were created to classify stars into spectral class, spectral subclass, and luminosity class. Since this tool is designed for accessible use, it was decided that instead of programming an entirely new Convolutional Neural Network (CNN), a type of machine learning model well-suited to image classification tasks, using Google's MobileNetV2 (Sandler et al., 2018; Sandler and Howard, 2018), a

Vol. 2024 (12) 560 – 570
ISSN 2688-3651 [online]

specialized type of CNN that was designed to be light and efficient, with applied transfer learning, where a previously trained model is trained on a new task to boost performance, would be a more efficient and low-cost method that could be more easily turned into a web application for other astronomers to use. As such, a machine-learning model was created in JavaScript using the Tensorflow.js (tfjs) library. The model itself was constructed of the MobileNetV2 with two added Dense layers, which helps put weights on different extracted features. The layers in the model filter down the image to figure out which features in the image are useful to classification and put weights on these features. In addition, HyperText Markup Language (HTML) and Cascading Style Sheets (CSS) were used to develop a frontend that connects the backend machine-learning model to an accessible user interface so that the web application could be accessed by professional and amateur astronomers alike. The structure of the machine learning model is shown in Figure 6.

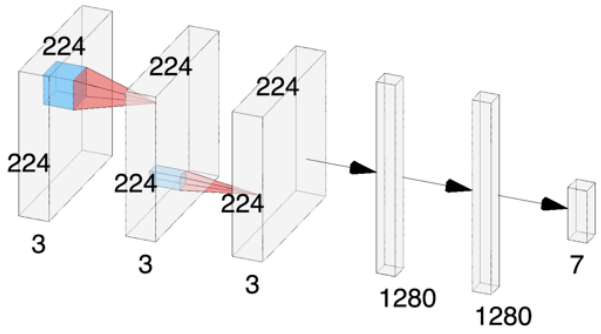


Figure 6. The structure of the machine learning model used to classify spectral class. 224 × 224 × 3 represents the input layer's dimensions, which corresponds with the image dimensions of 224 × 224 pixels, and three values in each pixel for red, green, and blue. The output layer has seven values, one for the machine learning model's confidence that an image falls in a class from O-M. The class with the highest confidence rating is selected as the output.

The machine learning model is designed to learn a bit like a human. Each time an image is fed into the machine learning model, it understands which parts of the image determine the classifications and which parts of the image to ignore, becoming more and more accurate each time it trains. The difference between such a model and a human is that the model is much more precise with the features in the image it uses to classify it.

In the machine learning model, the only parameter that was used the image input. For hyperparameters, after performing a grid search, which tries many different combinations to find the optimal one, it was decided to use a learning rate of 0.001 and a batch size of 16.

2.4 Spectroscopic Parallax Calculation from Classifications

After training the machine learning model to classify stars, the classifications were then used to calculate the distance to a star using the spectroscopic parallax method, which relies on comparing how bright a star actually is at a standardized distance, which is known as the absolute magnitude, to how bright the star looks from the Earth, known as the apparent magnitude, to find the distance. The first step in the spectroscopic parallax method was to calculate the absolute magnitude of the star. To do so, the output classifications from the machine learning model were used along with standard textbook data. An example of the data is shown in Table 3.

Table 3. An example of the standard textbook data to find the (visual) absolute magnitude, column MV, from the spectral type for luminosity class III (Carroll and Ostlie, 2018). Other tables show similar data for luminosity classes I and V.

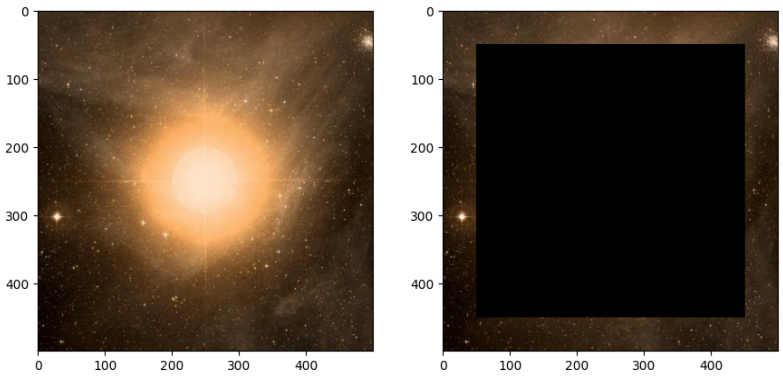
Giant Stars (Luminosity Class III)									
Sp. Type	$T_e\ (K)$	$L/L_\odot$	$R/R_\odot$	$M/M_\odot$	M_{bol}	BC	M_V	$U-B$	$B-V$
O5	39400	741000	18.5	—	−9.94	−4.05	−5.9	−1.18	−0.32
O6	37800	519000	16.8	—	−9.55	−3.80	−5.7	−1.17	−0.32
O7	36500	375000	15.4	—	−9.20	−3.58	−5.6	−1.14	−0.32
O8	35000	27700	14.3	—	−8.87	−3.39	−5.5	−1.13	−0.31

The second step in performing spectroscopic parallax was calculating the apparent magnitude from the image of the star. To do this, first, the pixel brightness of the entire image was found, and then the pixel brightness of the background of the image was subtracted to find the brightness of only the star. This brightness was then divided by

Vol. 2024 (12) 560 – 570
ISSN 2688-3651 [online]

that of a reference star, which in this case was Vega, to get the apparent magnitude. This process is shown in Figure 7.

The third and final step of performing spectroscopic parallax involves using distance modulus to compare the absolute and apparent magnitude values to find the distance to the star. The distance modulus formula is shown in Equation (3), where m is the apparent magnitude, M is the absolute magnitude, and d_{pc} is the distance to the star in pc (Carroll and Ostlie, 2018).



Original Image *Image with only the background*

Figure 7. An example of the output of background subtraction on an image using a specified cutout border-radius.

$$m - M = -5 + 5\,log_{10}\,d_{pc} \tag{3}$$

To directly get a value of d_{pc}, the equation was solved to yield Equation

$$d_{pc} = 10^{(m-M+5)/5} \tag{4}$$

Subsequently, for each star, the absolute and apparent magnitude were calculated, and then Equation 4 was applied to calculate the distance.

3. Results

3.1 Discussion of Error Metrics

Three error metrics were used in this section to determine the error in absolute magnitude, apparent magnitude, and distance. For absolute and apparent magnitude, relative root mean square error (RRMSE) was calculated as shown in Equation (4). For the distance, the mean percent error (MPE) as calculated in Equation (5) instead of RRMSE. In both of these equations, x_i is an actual value, $\hat{x}_i$ is a calculated/observed value, and N is the number of values in the set (which in this case is 61). In this case, MPE was used instead because it not only shows the average percentage error but is also a useful indicator of possible bias in a model. However, it was replaced by RRMSE for absolute and apparent magnitude as MPE fails where values are 0, and because some of the absolute and apparent magnitude values were 0, MPE was replaced with another error metric. Both metrics serve for useful comparisons as they are expressed as a percentage, meaning that they can be compared relative to one another because they are not raw data values.

$$RRMSE = \sqrt{\frac{\frac{1}{N}\sum_{i=1}^{N}(x_i - \hat{x}_i)^2}{\sum_{i=1}^{N}(\hat{x}_i)^2}} \tag{5}$$

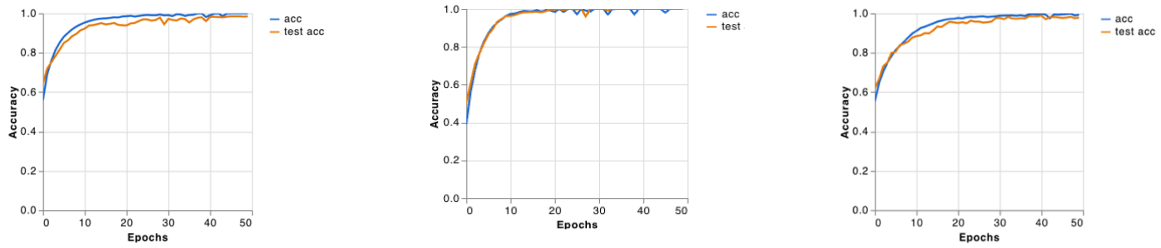
$$MPE = 100\% \times \frac{1}{N}\sum_{i=1}^{N}\frac{x_i - \hat{x}_i}{x_i} \tag{6}$$

3.2 Machine Learning Model Accuracy

Three machine-learning models reached 97.9%, 99.4%, and 97.9% accuracy, respectively, across classifications in spectral class, spectral subclass, and luminosity class. These machine learning models, constructed of a

Vol. 2024 (12) 560 – 570
ISSN 2688-3651 [online]

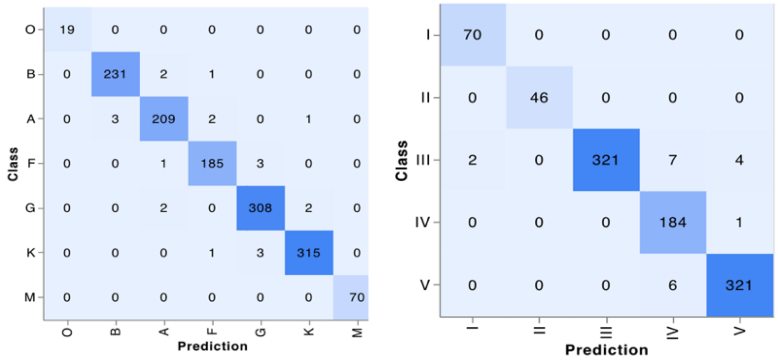
MobileNetV2 with applied transfer learning, were trained over 50 epochs with a learning rate of 0.001 and a batch size of 16. The training/testing accuracy graphs over the 50 epochs for each of the three machine-learning models are shown in Figure 8.



Spectral Class (97.9%) Spectral Subclass (99.4%) Luminosity Class (97.9%)

Figure8. Accuracy graphs over 50 epochs in training for the three machine learning models designed to identify the spectral type of a star. The blue line represents the training accuracy, and the orange line is the testing accuracy.

Upon manually testing the classifications with other images found of stars, it was found that the machine learning model was able to classify them all either completely correctly or with a very close classification that was off by a spectral subclass. The confusion matrix, which shows how the machine learning model classified the images, is shown in Figure 9.



Confusion Matrix for Spectral Class Confusion Matrix for Luminosity Class

Figure 9. Confusion matrices for the spectral and luminosity class. Items the machine learning model classified correctly go down a diagonal from the top left to the bottom right.

3.3 Absolute Magnitude Calculations

Given the spectral and luminosity classifications, the absolute magnitude can be calculated using the Hertzsprung-Russell Diagram data. In order to test the accuracy of the calculated absolute magnitude values, the calculated values were plotted against the verified, actual values scraped from various sources (Hummel et al., 2013). Perfect accuracy would mean that the plotted points form a straight, diagonal line going from the bottom left to the top right of the plot. The plot is shown in Figure 10.

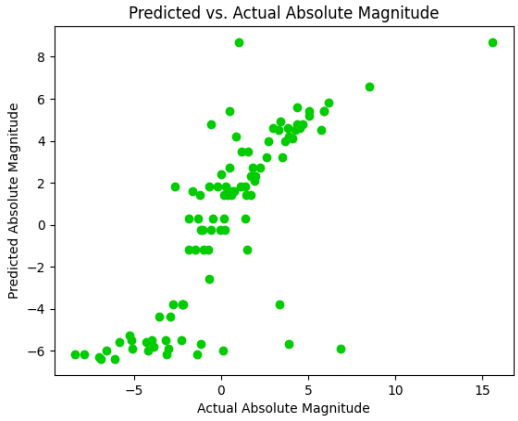


Figure 10. Calculated vs. Actual Absolute Magnitude for 61 stars with known absolute magnitudes that could be scraped from various online databases.

An r^2 value of 0.818 was calculated for the plotted points in **Error! Reference source not found.**. In addition to the r^2 value, the RRMSE regression error metric was also calculated, as shown in Equation 3, to understand the relative error in the calculations of absolute magnitude. Using such metrics allows for the relative comparison of error by percentage instead of by raw data value. RRMSE was calculated to be 6.59% for the absolute magnitude.

3.4 Apparent Magnitude Calculations

A similar procedure was conducted for the apparent magnitude as was done for the absolute magnitude. Calculated apparent magnitudes were plotted against the actual apparent magnitudes, as shown in Figure .

Vol. 2024 (12) 560 – 570
ISSN 2688-3651 [online]

An r^2 value of 0.651 was calculated for this plot. RRMSE was once again calculated, with a value of 37.0%, respectively. The error is relatively higher than that of the absolute magnitude, consistent with Figure 10 and Figure 11.

3.4 Distance Calculations

The calculated distance values from Equation (4) were plotted against the collected measured distance values from SIMBAD; the resulting graph is shown in Figure 12.

An r^2 value of 0.926 was calculated for this plot. The calculated MPE is -13.1%, indicating an underestimate bias.

In summary, the machine learning model is very accurate in classifying the spectral type, reaching near 100% accuracy. In addition, the calculations for the distances have a relatively high r^2 value of 0.926, indicating a relatively low error.

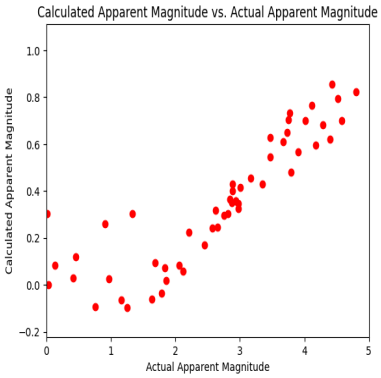


Figure 11. Calculated vs. Actual Apparent Magnitude for 61 stars with known apparent magnitudes that could be scraped from various online databases.

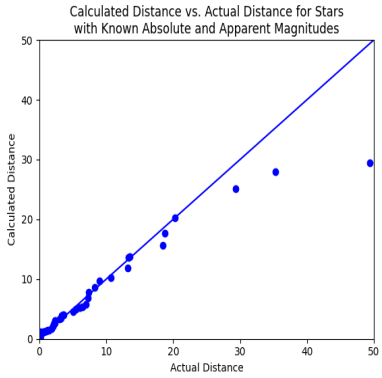


Figure 12. Calculated vs. Actual Apparent Magnitude for 61 stars with known apparent magnitudes that could be scraped from various online databases. The perfect-fit trendline is plotted for reference.

4. Discussion

4.1 Error Analysis

Error in the absolute magnitude mainly stems from the approximations made using Hertzsprung-Russell Diagram data. For such data, since absolute magnitude was determined using Table 3, where a specific spectral and luminosity class were mapped to an absolute magnitude, there is a certain amount of inaccuracies due to variation even for stars having the exact same spectral type. Since classification is used instead of regression, there will be a certain amount of error due to variation within a class.

The apparent magnitude has a notably higher error than the absolute magnitude. One possible source of such error is with the background subtraction method. Because of the difficulty of completely accurately removing the background from the image, for some images, there will be too little background removed, whereas, for others, there will be too much removed. Apparent magnitude is usually calculated using data from the charge-coupled devices (CCD) of a telescope; however, such data was unavailable for usage in this research project. In the future, utilizing such data in this research would create a more effective method for calculating the distance using spectroscopic parallax.

In addition, incorrect classifications could affect the accuracy of the distance calculation. Incorrectly classifying the spectral subclass would not have a significant effect on the distance calculation, but incorrectly classifying the spectral or luminosity class would have a much larger effect as that means the star is given an entirely different range of values for the distance estimate. As the accuracy of the classifications is near 100%, this error is marginalized in this research.

The distance calculation has a relatively low error compared to what may be expected by most of a method that relies only on an image of a star. However, it can be seen that, for larger distances, the calculated distance tends to start falling short of the actual distance. This is most likely due to the calculation of the apparent magnitude falling short of the actual; while it is accurate for a certain range, as the actual distance becomes larger, the difference becomes more noticeable. Another possible source of error here could also be due to different obstacles causing a star to look different than it actually does. For example, dust extinction, where electromagnetic radiation is absorbed and scattered by interstellar dust and gas, causes a star to look redder than it actually is, a phenomenon known as interstellar reddening. To account for this, redshift and blueshift may also have to be taken into account for future models. In

Vol. 2024 (12) 560 – 570
ISSN 2688-3651 [online]

addition, although atmospheric interference could affect the images taken by telescopes on Earth, most observatories utilize what is known as an adaptive optics system to neutralize atmospheric interference in images taken. Telescopes in space have no atmospheric interference.

4.2 Future Work

There are multiple improvements and extensions that can be made to this research in the future. One such improvement is improving upon the apparent magnitude calculation. As of now, the calculations for apparent magnitude rely on performing background subtraction on an image. However, a caveat with this method is the existing inaccuracy in removing a background from an image. To correct for this in the future, real telescope CCD data could be used. While much of this CCD data is, unfortunately, currently unavailable online, one possible way to obtain such data would be to work with other professional astronomers in the field, gather data from the existing telescopes, and incorporate that into this research. As it stands, the tool is still quite useful in classifying stars and finding the distance, but CCD data could add an improvement to the accuracy.

Another improvement that could be made would be to gather more Hertzsprung-Russell Diagram data. For this research, existing data was used; however, this existing data only covered luminosity classes I, III, and V, and data points for certain spectral classes were missing. In order to account for this, a new standardized set of data could be created that is able to cover a wider range of spectral and luminosity classes. This would yield better calculations for the absolute magnitude, and therefore help make the distance calculation more accurate.

In the future, one possibility for how this work could be used is by applying it to a telescope in real-time. As a telescope takes images of stars, each image could be passed into this model to get an initial estimate for the distance to the star. This would allow for rapid distance calculations in real time.

5. Conclusion

This research used three machine learning models to classify stars into the Morgan-Keenan System, and then subsequently used these classifications and Hertzsprung-Russell Diagram data to find the distance to these stars. While existing tools can classify stars, none were found that can both classify and find the distance to a star. Three main steps were followed throughout the process of this research: data collection, in which data was collected on 430 stars; machine-learning model training, in which three separate machine-learning models were trained to identify a star and classify it; and distance calculation, in which absolute and apparent magnitude were calculated, and then distance modulus was applied to find the distance. Both the absolute magnitude and distance calculations were relatively accurate and had relatively low error rates, while the apparent magnitude had a slightly higher error rate. Future improvements include the addition of more data for an improvement in accuracy as well as an extension of this tool to other celestial objects. The output of this research allows astronomers to take rapid data measurements on stars and also find the distance to these stars in real-time as photos are taken. Adding such a tool to today's telescopes will vastly improve the capabilities that these telescopes have, and will also make them more cost-effective in the long run.

Acknowledgment

I would like to thank my mentor for this research project, Dr. Mary Krasovec, who provided me with much support and valuable insights throughout the research process. In addition, I would like to thank Victoria Lloyd for their assistance with the technical and writing portions of this research.

References

Abadi, M., et al. (2016). TensorFlow: A system for large-scale machine learning. ArXiv:1605.08695 [Cs]. https://arxiv.org/abs/1605.08695

Vol. 2024 (12) 560 – 570
ISSN 2688-3651 [online]

Althaus, L., Córsico, A., Isern, J., & Garcia-Berro, E. (2010). Evolutionary and pulsational properties of white dwarf stars. *The Astronomy and Astrophysics Review*, 18, 471–566. ResearchGate. https://www.researchgate.net/publication/225677859_Evolutionary_and_pulsational_properties_of_white_dwarf_stars

Amos, J. (2016, September 14). Gaia space telescope plots a billion stars. BBC News. https://www.bbc.com/news/science-environment-37355154

Bradski, G. (2000). The OpenCV Library. Dr. Dobb's Journal of Software Tools.

Brau, J. (2016, October 16). Astronomy 122 - Measuring the Stars. Astronomy 122 - Measuring the Stars. https://pages.uoregon.edu/jimbrau/astr122/Notes/Chapter17.html

Budassi, P. C. (2024). A simple chart for classifying the main star types using Harvard classification. Wikimedia.org. https://upload.wikimedia.org/wikipedia/commons/3/37/Stellar_Classification_Chart.png

Buslaev, A., et al. (2020). Albumentations: Fast and Flexible Image Augmentations. Information, 11(2), 125. https://doi.org/10.3390/info11020125

Carroll, B. W., & Ostlie, D. A. (2018). An introduction to modern astrophysics (2nd ed.). Cambridge University Press.

Chandra, R. V., & Varanasi, B. S. (2015). Python requests essentials. Packt Publishing Ltd.

Clark, A. (2015). Pillow (PIL Fork) Documentation. readthedocs. Retrieved from https://buildmedia.readthedocs.org/media/pdf/pillow/latest/pillow.pdf

Dafonte, C., et al. (2020). A Blended Artificial Intelligence Approach for Spectral Classification of Stars in Massive Astronomical Surveys. *Entropy*, 22(5), 518. https://doi.org/10.3390/e22050518

Despotovic, M., et al. (2016). Evaluation of empirical models for predicting monthly mean horizontal diffuse solar radiation. *Renewable and Sustainable Energy Reviews*, 56, 246–260. https://doi.org/10.1016/j.rser.2015.11.058

Ginsburg, A., et al. (2019). astroquery: An Astronomical Web-querying Package in Python. *The Astronomical Journal*, 157(3), 98. https://doi.org/10.3847/1538-3881/aafc33

Harris, C. R., et al. (2020). Array programming with NumPy. Nature, 585, 357–362. https://doi.org/10.1038/s41586-020-2649-2

hips2fits - Fast generation of FITS cutouts from HiPS datasets. (n.d.). Alasky.cds.unistra.fr. Retrieved March 25, 2024, from https://alasky.cds.unistra.fr/hips-image-services/hips2fits

Hummel, C. A., et al. (2013). Dynamical mass of the O-type supergiant in Zeta Orionis A. *Astronomy & Astrophysics*, 554, A52. https://doi.org/10.1051/0004-6361/201321434

Hungund, G. (2020). Computational Astronomy: Classification of Celestial Spectra Computational Astronomy: Classification of Celestial Spectra Using Machine Learning Techniques Using Machine Learning Techniques Part of the Artificial Intelligence and Robotics Commons, External Galaxies Commons, and the Other Astrophysics and Astronomy Commons. https://core.ac.uk/download/pdf/323446272.pdf

International Astronomical Union | IAU. (2015). Iau.org. https://www.iau.org/public/themes/naming_stars/
Kalhor, B. (2023). Data of nearby space objects using SIMBAD astronomical database. *Data in Brief*, 47(47), 108943. https://doi.org/10.1016/j.dib.2023.108943

Kalinin, A. A. (2018). Albumentations: fast and flexible image augmentations. ArXiv E-Prints.

Vol. 2024 (12) 560 – 570
ISSN 2688-3651 [online]

Kuntzer, T., Tewes, M., & Courbin, F. (2016). Stellar classification from single-band imaging using machine learning. *Astronomy & Astrophysics*, 591(A&A), A54. https://doi.org/10.1051/0004-6361/201628660

Kyritsis, E., et al. (2022). A new automated tool for the spectral classification of OB stars. *Astronomy & Astrophysics*, 657, A62. https://doi.org/10.1051/0004-6361/202040224

Langendorf, R., Schneider, S., & Klein, P. (2022). Extracting information from the Hertzsprung-Russell diagram: An eye-tracking study. *Physical Review Physics Education Research*, 18(2). https://doi.org/10.1103/physrevphyseducres.18.020121

Langer, N., & Kudritzki, R. P. (2014). The spectroscopic Hertzsprung-Russell diagram. *Astronomy & Astrophysics*, 564(A&A), A52. https://doi.org/10.1051/0004-6361/201423374

Magee, J. (2021). Stellar classification of folded spectra using the MK Classification Stellar classification of folded spectra using the MK Classification scheme and convolutional neural networks scheme and convolutional neural networks. *Arrow@TU Dublin*. https://doi.org/10.21427/SETB-YD94

Morgan, W., Keenan, P., & Kellman, E. (1943). An Atlas of Stellar Spectra with an Outline of Spectral Classification. In *Astrophysical Monographs*. The University of Chicago Press. https://www.ucl.ac.uk/ucl-observatory/sites/ucl_observatory/files/mkkbook.pdf

Perez, L., & Wang, J. (2017). The Effectiveness of Data Augmentation in Image Classification using Deep Learning. ArXiv:1712.04621 [Cs]. https://arxiv.org/abs/1712.04621

Price-Whelan, A. M., et al. (2018). The Astropy Project: Building an open-science project and status of the v2. 0 core package. *The Astronomical Journal*, 156(3), 123.

Qi, Z. (2022). Stellar Classification by Machine Learning. *SHS Web of Conferences*, 144, 03006. https://doi.org/10.1051/shsconf/202214403006

Roulston, B. R., Green, P., & Kesseli, A. Y. (2020). Classifying Single Stars and Spectroscopic Binaries Using Optical Stellar Templates. *The Astrophysical Journal Supplement Series*, 249(2), 34–34. https://doi.org/10.3847/1538-4365/aba1e7

Sandler, M., & Howard, A. (2018, April 3). MobileNetV2: The Next Generation of On-Device Computer Vision Networks. Blog.research.google. https://blog.research.google/2018/04/mobilenetv2-next-generation-of-on.html

Sandler, M., et al. (2018). MobileNetV2: Inverted Residuals and Linear Bottlenecks. ArXiv.org. https://arxiv.org/abs/1801.04381

Shi, J.-H., et al. (2023). Stellar classification with convolutional neural networks and photometric images: a new catalogue of 50 million SDSS stars without spectra. *Monthly Notices of the Royal Astronomical Society*, 520(2), 2269–2280. https://doi.org/10.1093/mnras/stad255

Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on Image Data Augmentation for Deep Learning. *Journal of Big Data*, *6*(1). https://doi.org/10.1186/s40537-019-0197-0

Xiao-Qing, W., & Jin-Meng, Y. (2021). Classification of star/galaxy/QSO and star spectral types from LAMOST data release 5 with machine learning approaches. *Chinese Journal of Physics*, 69, 303–311. https://doi.org/10.1016/j.cjph.2020.03.008

Zhao, Z., Wei, J., & Jiang, B. (2022). Automated Stellar Spectra Classification with Ensemble Convolutional Neural Network. *Advances in Astronomy*, 2022, 1–7. https://doi.org/10.1155/2022/4489359