A Hybrid Photometric and Spectral Algorithm for Efficient Detection of Gravitationally Lensed Quasars

Pranav R. Sivakumar, Janani N. Sivakumar

BRIEF. A novel method was developed to identify gravitationally lensed quasars, consisting of a morphological algorithm directed at finding wide-separation lens candidates and a point spread function (PSF)-difference-based algorithm aimed at identifying close-separation lens candidates.

ABSTRACT. A novel method was developed to identify gravitationally lensed quasars from the Sloan Digital Sky Survey (SDSS). The method consisted of two algorithms: a morphological algorithm directed at finding wide-separation lens candidates and a point spread function (PSF)-difference-based algorithm aimed at identifying close-separation lens candidates. Understanding gravitational lensing can help decipher the properties of dark matter and dark energy. It is hypothesized that if multiple objects in an SDSS image meet both spectral and photometric criteria, then these objects are potentially images of the same gravitationally lensed quasar. This project compiled data from over 300,000 quasars in the SDSS Data Release 10 and 592,313 neighbors within 16 arcseconds of each quasar. The data was retrieved and processed using Structured Query Language queries. The algorithms compared the quasars to their neighbors to determine if the neighbors were images of the same quasar. The results were validated against a control group of lensed quasars reported in the literature. A comparison of the project's results with established data sets of lensed quasars led to the conclusion that the hypothesis was well supported. In addition to identifying a majority of the quasars in the control group, the algorithms also identified new high-probability lens candidates not yet reported in the literature.

INTRODUCTION.

Quasars, also known as QSOs (quasi-stellar objects), are point-like luminous astronomical objects located around black holes. QSOs are some of the oldest and farthest objects known. Since they are very distant, the electromagnetic radiation coming from them can be affected by the masses of their host galaxies, as well as other objects between Earth and the quasars.

General relativity predicts that the gravity of a massive object interposed between a light source and an observer bends the light from the source. This phenomenon, known as gravitational lensing, can make the source appear brighter and produce multiple images of the source. Gravitational lensing of quasars, classified as strong lensing, is particularly valuable; it furthers the understanding of dark matter and dark energy, which constitute about 95 percent of the universe but are poorly understood [1].

As the phenomenon of lensing is caused solely by gravity, it is the only way to directly determine the distribution of dark matter in the universe. The results of lensing studies not only give scientists a better understanding of dark matter, but can also provide information about the amount of dark energy in the universe. The balance of dark matter and dark energy determines whether the universe will continue expanding forever, stay in its current state, or collapse [1, 2].

Despite the utility of quasar lensing, only a small number of lensed quasars have been discovered to date. Since 1979, when the first lensed quasar was discovered, only 120 lensed quasars have been confirmed. With Data Release 10 (DR10) of the SDSS, spectral and photometric data became available for over 300,000 quasars; this is among the largest quasar data sets to date [3]. The scientific value of studying lensed quasars and the availability of extensive data from DR10 motivated the authors to pursue this research project.

The significance of quasar lensing in solving larger cosmological puzzles has made it a very active area of research. A pioneering 1984 paper [4] launched the study of quasar lensing by examining the likelihood that a quasar will be lensed by a galaxy or by a point mass. The next major study of quasar lensing was the Hubble Space Telescope Snapshot Survey [5], which showed that about one percent of bright quasars at z > 1 (z = redshift) are gravitationally lensed into two or more images.

The search for lensed quasars accelerated with the arrival of SDSS DR3 data in 2005. Masamune Oguri and others presented an algorithm to identify gravitationally lensed quasars that formed the basis of the SDSS Quasar Lens Search (SQLS) [6]. The algorithm used a photometric method to identify close-separation lensed quasars and color selection techniques to identify lenses with larger separations.

Following up their earlier work, the SQLS team applied quasar lensing probabilities to establish limits on the value of the cosmological constant and the distribution of dark energy. By comparing the observed lensing fraction with theoretical models, they determined that their results support the accelerated expansion model of the universe [2].

Prior to the present research work, quasar lensing research focused solely on either photometric or spectral characteristics of quasars. The candidate selection algorithms developed in this work combine both types of data to identify lensed quasar candidates, thus improving the accuracy and reliability of the candidates identified for follow-up observations.

MATERIALS AND METHODS.

This research uses morphological and point spread function (PSF)-difference algorithms to find candidate lists of lensed quasars from the SDSS DR10. Working in tandem, the two algorithms can identify lensed quasar candidates over a wide range of image separations. The candidates from this project are validated against lensed quasars from the Master Lens Database [7], one of the largest compilations of known and verified lensed quasars.

Morphological Algorithm

The morphological algorithm first examines all objects identified as quasars by the SDSS pipeline. The pipeline is able to classify most quasars, but it sometimes misses quasars spread over a large area. In order to fill this gap, a query using color selection criteria from the literature [8] was created to select objects not identified as quasars by the SDSS. The standard quasar and color cut data sets were combined and duplicates were removed, resulting in a baseline data set of 532,704 quasar candidates. The addition of objects with quasar-like color characteristics added almost 200,000 quasar candidates to the data set.

For each quasar in the baseline data set, all the neighbors within a separation threshold were identified. Based on analysis of published literature, a separation value of 16 arcseconds (") was selected. There were 592,313 neighbors located within 16" of the target quasars.

Spectral data was retrieved from the SDSS SpecPhotoAll table, which includes all the spectra associated with the quasars and neighbors. Some neighbors of the target quasars lacked spectral data. If the neighbor had spectral data, its redshift was compared to that of the target quasar. The redshift difference criterion was set as a percentage of the target quasar redshift in order to make it adaptive.

The SDSS captures images through five filters – ultraviolet (u), green (g), red (r), infrared (i), and far infrared (z). Astronomical color is determined by taking

the difference between the magnitudes of an object measured through two different color filters such as g and r. The g-r color of each neighbor was compared to the g-r color of the target quasar. If a neighboring object has a similar redshift and color to the target quasar, it is likely that the target quasar and the neighbor are two images of the same lensed quasar. If the neighbor had no spectral data, only the color comparison was utilized.

The color comparison was performed using the functions below:

$$D(i-j)=(i-j)_{quasar}-(i-j)_{neighbor}$$

$$OR$$

$$D(i-j)=(i-j)_{neighbor}-(i-j)_{quasar}$$

where $D(i-j)$ is the color difference between the quasar and its neighbor in a particular pair of color bands. The absolute value of $D(i-j)$ was not used to account for the fact that the value of $(i-j)_{neighbor}$ could be positive or negative. The first equation was used if this value was positive; otherwise, the second equation was utilized in the color comparison.

Candidates which met the necessary conditions of the algorithm were classified into three types. Type 1 candidates matched both redshift and color criteria, while Type 2 candidates matched only the redshift criterion. The candidate quasar-neighbor pairs classified as Type 3 did not have spectral data and hence were selected using only the color criterion.

As a final validation step, the spectra of the high-probability candidates were visually compared against the spectra of the neighbors. A template similar to the layout shown in Figure S1 was used for visual review. The visual analysis compared the locations and other features of the emission lines in the two spectra (annotated in Figure S1) as well as their overall shape. If the spectra exhibited similar characteristics, the candidate was placed in the final high-probability candidate list.

PSF-Difference Algorithm

The morphological algorithm is unable to identify close-separation lenses not deblended (identified as two objects) by the SDSS pipeline. In this research, PSF-based criteria were used to deblend close quasar-neighbor pairs.

The SDSS calculates a PSF magnitude (psfMag) by fitting a model PSF to an object. When a PSF is fitted to an object that SDSS cannot deblend, the PSF magnitude is significantly different from the normal SDSS magnitude (model-Mag). Based on analysis of known lensed quasars, the minimum and maximum PSF difference criteria were set at 0.13 and 0.4, respectively.

Images of the candidates that met the PSF difference criteria were placed into a template similar to the layout shown in Figure S1 for visual examination. Subsequently, the candidates were classified into types 1, 2, and 3 based on the visual similarity of the images.

The candidates were then cross-matched against three sky surveys in different bands of the electromagnetic spectrum: the Faint Images of the Radio Sky at Twenty-Centimeters (FIRST) survey; the Röntgensatellit (ROSAT) All-Sky Survey in X-ray bands; and the AllWISE survey, carried out using the Wide-Field Infrared Survey Explorer (WISE). Successful cross-matches provide a higher degree of confidence. However, non-matches do not rule out the possibility of quasars being lensed.

The overall candidate selection algorithm is outlined in Figure 1.

Data Extraction and Processing

The first steps of data analysis were data extraction and synthesis, criterion matching, and data validation. The baseline set of quasars for the morphological algorithm was compiled using an SQL query which returned the spectral and photometric parameters required for the different comparison steps. SQL queries were also used to extract data and set criteria for the PSF-difference algorithm. The resulting data sets were analyzed and visualized using scientific packages in Python. Statistical analyses of input data sets and results were performed using R and JMP.
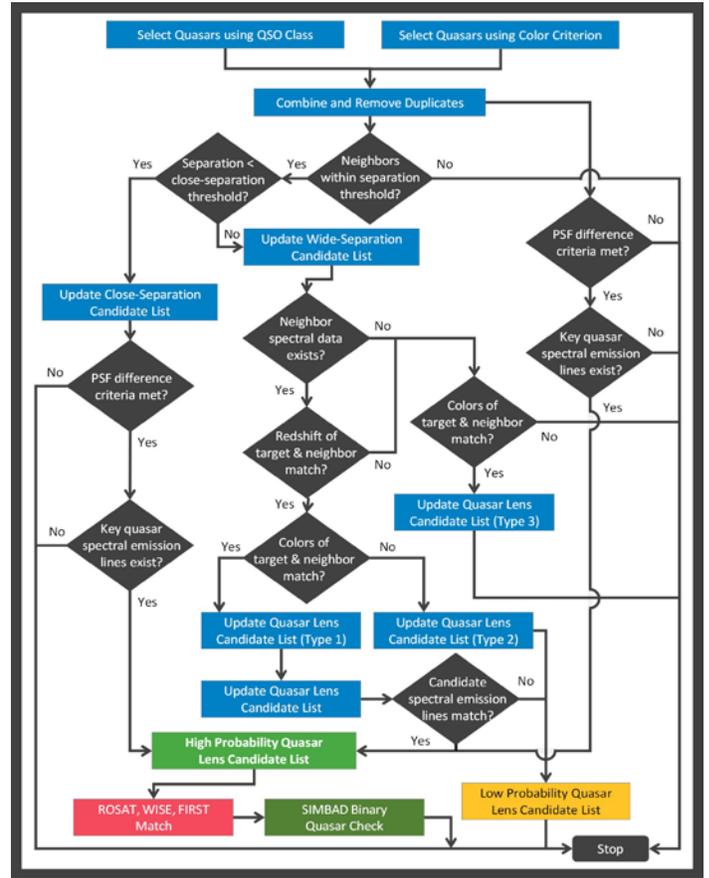
Figure 1. Flowchart of the candidate selection algorithms.

RESULTS.

The selection criteria and number of preliminary candidates identified by the morphological algorithm are listed in Table S1. The use of restrictive redshift and color criteria resulted in fewer successful matches in Types 1-2 as compared to Type 3. The Type 1 candidates were further organized into four subtypes based on the SDSS classes of the targets and the neighbors. Quasar-quasar lens candidates were classified as Type 1A, quasar-galaxy pairs as Type 1B, galaxy-quasar pairs as Type 1C, and galaxy-galaxy pairs as Type 1D. The Type 2 candidates were classified according to the same scheme.

The spectral comparison of 278 Type 1 candidates produced 42 final candidates, as shown in Table 1. There were 34 Type 1A candidates, 7 Type 1B candidates, and 1 Type 1D candidate. The fact that the algorithm identified only 42 final candidates out of almost 600,000 quasar-neighbor pairs illustrates the rarity of quasar lensing. The relative lack of Type 1D candidates was due to the extended-source nature of galaxies as opposed to the point-source nature of quasars.

The PSF-difference algorithm successfully identified a number of lensed quasars that could not be detected using the morphological algorithm. The initial PSF difference query selected 17,488 SDSS candidates. After visual examination, 146 candidates remained – 57 Type 1 candidates, 63 Type 2 candidates, and 26 Type 3 candidates. Table 1 lists the data coverage, selection criteria, and number of preliminary candidates identified. The PSF-difference algorithm not only identified a number of lensed quasars from the literature, but also 93 new lensed quasar candidates hitherto unreported in the literature.

Statistical Analysis

A set of statistical analyses was performed on both the input data and the resulting candidate lists. The candidate quasars were binned into one of eight groups based on their redshift values, and a Student's t-test was performed to compare the mean value of separation in each bin. The analysis results, shown in Figure

S2, confirmed that the spatial separation threshold is consistent across the entire redshift range.
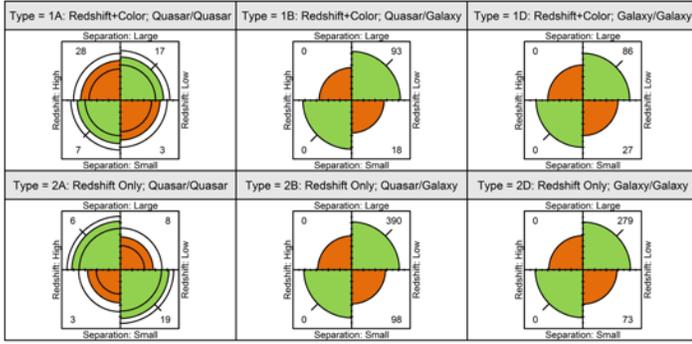
Figure 2. Categorical analysis of separation against redshift stratified by candidate type.

Figure 2 shows the results from a categorical analysis of separation against redshift using candidate type as the stratum (Type 1C and 2C candidates were not included due to the lack of a sufficient number of candidates). The candidates were grouped into two redshift categories and two separation categories: high-redshift ($z > 2$), low-redshift ($z \leq 2$), large-separation (> 5"), and small-separation (≤ 5").

The odds ratio indicates the strength of the association between redshift and separation and answers questions such as "Do high-redshift quasars tend to have larger separations than low-redshift quasars?". There was no statistically significant association for Type 1A, as the quadrants roughly formed a circle, indicating an odds ratio of about 1. However, all the other candidate subtypes displayed a statistically significant relationship between redshift and separation.

An outlier analysis, shown in Figure S3, was also performed to determine how well the parameters of each high-probability candidate were grouped together. The analysis indicates that the redshift differences of some quasar-quasar pairs are outliers. However, the outliers were not removed from the data, since the differences were still within the redshift threshold.

Table 1. Summary of candidates identified by the morphological and PSF-difference algorithms.

High-probability candidates detected by morphological algorithm			
Candidate List	Scope	Criteria	Number of Candidates
Type 1A	Spectral and Photometric Data	Quasar.Class = 'QSO' and Neighbor.Class = 'QSO' Redshift and Color	34
Type 1B	Spectral and Photometric Data	Quasar.Class = 'QSO' and Neighbor.Class = 'Galaxy' Redshift and Color	7
Type 1D	Spectral and Photometric Data	Quasar.Class = 'Galaxy' and Neighbor.Class = 'Galaxy' Redshift and Color	1

High-probability candidates detected by PSF-difference algorithm			
Candidate List	Visual Observation Criteria	Number of Candidates Reported in Literature	Number of Candidates Not Yet Reported in Literature
Type 1	High level of similarity	29	28
Type 2	Medium level of similarity	16	47
Type 3	Low level of similarity	8	18

DISCUSSION.

The lensed quasars identified by the morphological and PSF-difference approaches were compared to the control group from the Master Lens Database. However, it is not appropriate to use all 120 quasar lenses in the Master Lens Database for comparison. One group of lensed quasars in the database fell outside the sky coverage of the SDSS, while another set of lenses had very low magnitudes. After the unmatchable lenses were removed, there were 34 lensed quasars that were relevant for comparison with the results of the morphological algorithm. The algorithm was able to match 88% of the relevant candidates in the control group.
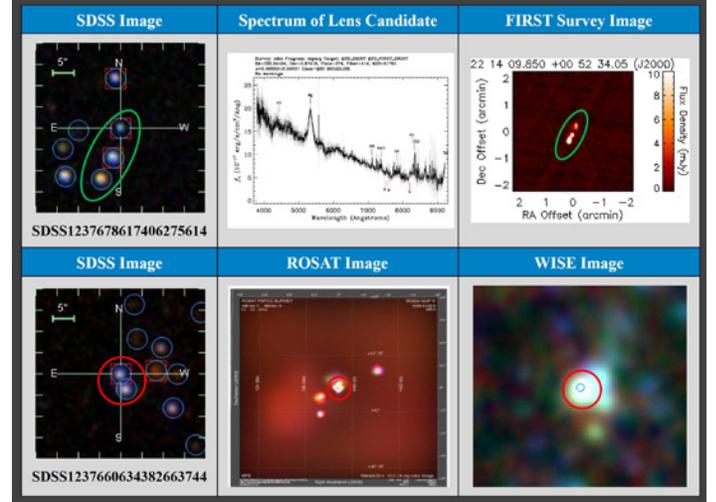
Figure 3. A representative sample of morphologically identified lensed quasar candidates with matching FIRST, ROSAT, and WISE images. Green ovals show corresponding objects in SDSS and FIRST images. Red circles highlight corresponding objects in SDSS, ROSAT, and WISE images.

As previously discussed, one of the last steps of the algorithm compared the SDSS images to FIRST, ROSAT and WISE images. Figure 3 shows two candidates identified by the algorithm; one candidate is shown with the spectrum of the quasar and the corresponding FIRST image, and the second candidate is shown with its corresponding ROSAT and WISE images. Just as in the SDSS image, the quasar and its possible lensed image are visible in the FIRST, ROSAT, and WISE images. The objects do not appear identical in all of the images because quasars emit light differently in different bands of the electromagnetic spectrum.

The morphological approach, in addition to matching several known lensed quasars, also found a number of new high-probability candidates. Figure S1 shows a lens candidate identified by the morphological algorithm that has not been reported in the literature. A comparison of the spectrum of the target quasar with that of the lens candidate in the third column shows that the spectra have identical characteristics.

Often, binary quasars are mistakenly identified as lensed quasars, as the two quasars in a binary system can have very similar characteristics. The most famous case of confusion between binary and lensed quasars is that of QSO B2345+007A and QSO B2345+007B [9]. Although these quasars have very similar optical images and spectra, observations by the Chandra X-ray Telescope have shown that they are binary quasars.

In order to avoid cases such as QSO B2345+007, the lens candidates were verified against known binary quasars from the literature. The right ascension and declination of each candidate (equivalent to latitude and longitude on the Earth) were queried in the SIMBAD astronomical database against known catalogs of binary quasars. Of the 42 high-probability Type 1 candidates that met all the criteria of the morphological algorithm, six candidates have been identified as binary quasars and thus eliminated from consideration as lens candidates.

The PSF-difference algorithm effectively complements the morphological algorithm; together, they successfully identified a number of candidates. Of the 24 candidates identified as one photometric object by the SDSS, nine were matched by the PSF-difference algorithm. Six of the candidates identified by the PSF-difference algorithm are shown in Figure S4. The first row contains candidates that matched lenses in the control group. The candidates in the second row have not yet been reported in the literature.

In summary, the approach of using both the spectral and photometric characteristics of quasars to identify lens candidates distinguishes this research from similar work.

CONCLUSION.

The morphological and PSF-difference approaches are both effective in identifying lensed quasar candidates. The morphological algorithm identified many new wide-separation candidates, while the PSF difference algorithm found a number of new close-separation candidates. Follow-up observations of high-probability candidates using large telescopes are necessary to confirm the lensing effects.

The most logical future extension to this project is gravitational lens modeling. Lens modeling will help identify potential lensing objects; if there is no visible lensing object, then lensing by dark matter is a possibility.

Another area for future research involves more rigorous cross-matching with surveys in other bands of the electromagnetic spectrum. This is of crucial importance in differentiating quasars that are truly lensed from binary quasars.

The accuracy of the algorithm could be improved by using a complete set of spectral data, more detailed statistical analyses to refine threshold values, and implementation of additional validation steps in the algorithm. Opportunities to apply the results of this work to constrain cosmological parameters through lensing statistics (distribution of redshifts and separations) are currently being explored.

SUPPORTING INFORMATION.
Figure S1. A morphologically identified candidate not yet reported in the literature.
Figure S2. Analysis of spatial separation vs. redshift for all candidate quasars.
Table S1. Candidates matching different candidate selection criteria.
Figure S3. Outlier analysis of spatial separation for Type 1 candidates.
Figure S4. Lensed quasar candidates identified by the PSF-difference algorithm.
Acknowledgment S1. Official SDSS and SIMBAD acknowledgement statements.

REFERENCES.
1. Frieman, J.,et al., *Ann. Rev. Astron. Astrophys*, 46, 385-432 (2008).
2. Oguri, M., et al., *Astronomical J.*, 143, 120-134 (2012).
3. Ahn, C. P., et al., *ApJS*, 211.17 (2014).
4. Turner, E.L., et al., *ApJ*, 284, 1-22 (1984).
5. Maoz, D., et al., *ApJ*, 409, 28-41 (1993).
6. Oguri, M., et al., *Astronomical J.*, 132, 999-1013 (2006).
7. Moustakas, L. A., et al., *American Astronomical Society, AAS Meeting*, #219, #146.01 (2012).
8. Richards, G. T., et al., *Astronomical J.*, 121, 2308-2330 (2001).
9. Green, P., et al., *ApJ*, 571, 721-732 (2002).

Pranav Sivakumar is a sophomore and Janani Sivakumar is a senior at the Illinois Mathematics and Science Academy, Aurora, Illinois. They performed their research at home.