Quasar Island – Three new $z \sim 6$ quasars, including a lensed candidate, identified with contrastive learning

Xander Byrne [1,2]⋆, Romain A. Meyer [2,3], Emanuele Paolo Farina [4], Eduardo Bañados [2], Fabian Walter [2], Roberto Decarli [5], Silvia Belladitta [2,5], Federica Loiacono [5]

[1]*Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge, CB3 0HA, UK*
[2]*Max Planck Institut für Astronomie, Königstuhl 17, D-69117, Heidelberg, Germany*
[3]*Department of Astronomy, University of Geneva, Chemin Pegasi 51, 1290 Versoix, Switzerland*
[4]*Gemini Observatory, NSF's NOIRLAB, 670 N A'ohoku Place, Hilo, HI 96720, USA*
[5]*INAF – Osservatorio di Astrofisica e Scienza dello Spazio di Bologna, Via Gobetti 93/3, 40129 Bologna, Italy*

ABSTRACT

Of the hundreds of $z > 6$ quasars discovered to date, only one is known to be gravitationally lensed, despite the high lensing optical depth expected at $z \gtrsim 6$. High-redshift quasars are typically identified in large-scale surveys by applying strict photometric selection criteria, in particular by imposing non-detections in bands blueward of the Lyman-α line. Such procedures by design prohibit the discovery of lensed quasars, as the lensing foreground galaxy contaminates the photometry of the quasar. We present a novel quasar selection methodology, applying contrastive learning (an unsupervised machine learning technique) to Dark Energy Survey imaging data. We describe the training of a neural network which isolates an 'island' of 11 sources, of which 7 are known $z \sim 6$ quasars. Of the remaining four, three are newly discovered quasars (J0109–5424, $z = 6.07$; J0122–4609, $z = 5.99$; J0603–3923, $z = 5.94$), as confirmed by follow-up Gemini-South/GMOS and archival NTT/EFOSC2 spectroscopy, implying a 91% efficiency for our novel selection method; the final object on the island is a brown dwarf. In one case (J0109–5424), emission below the Lyman limit unambiguously indicates the presence of a foreground source, though high-resolution optical/near-infrared imaging is still needed to confirm its lensed (multiply-imaged) nature. Detection in the g band has led this quasar to escape selection by traditional colour cuts. Our findings demonstrate that machine learning techniques can thus play a key role in unveiling populations of quasars missed by traditional methods.

Key words: quasars: individual (J0109–5424, J0122–4609, J0603–3923) – quasars: supermassive black holes – gravitational lensing: strong

1 INTRODUCTION

High-redshift ($z \gtrsim 6$) quasars are crucial probes of the early Universe, constraining the early growth of supermassive black holes, the properties of the intergalactic medium (IGM), and likely tracing the formation of the first massive structures in the Universe (see e.g. Inayoshi et al. 2020; Volonteri et al. 2021; Fan et al. 2023 for reviews).

Multi-wavelength large-sky surveys, such as the Dark Energy Survey (DES; Flaugher et al. 2015), the VISTA Hemisphere Survey (VHS; McMahon et al. 2013), the Wide-field Infrared Survey Explorer (WISE; Wright et al. 2010), the Panoramic Survey Telescope and Rapid Response System (Pan-STARRS1; Chambers et al. 2016) and the Sloan Digital Sky Survey (SDSS; York et al. 2000), have enabled the discovery of several hundred high-redshift quasars to date. Quasars are typically distinguished from other sources, such as stars and galaxies, by their distinctive colour (e.g., Venemans et al. 2013; Bañados et al. 2015; Reed et al. 2015, 2017). The high opacity of the neutral IGM at $z > 5.5$ creates a distinctive break below the Lyman-α

line, and no flux is expected to be transmitted blueward of the rest-frame Lyman limit ($\lambda = 912$ Å; redshifted to $\lambda > 6500$Å). Hence traditional $z \gtrsim 6$ quasar dropout selection criteria often impose non-detections in the g and r bands (e.g., Fan et al. 2001; Richards et al. 2002; Reed et al. 2015; Jiang et al. 2016; Bañados et al. 2016; Reed et al. 2017; Wang et al. 2019; Bañados et al. 2023).

Such selection criteria are not perfect, and often include contaminants. These are most often Galactic brown dwarfs, which have similar colours to high-redshift quasars and are thus difficult to systematically remove (e.g., Wang et al. 2019; Yang et al. 2022). Moving solar system objects can also contaminate high-redshift quasar searches (Bosman et al. 2023). Removing these contaminants requires time-consuming refinement procedures such as visual inspection (e.g., Venemans et al. 2013; Bañados et al. 2015; Reed et al. 2015), SED fitting (e.g., Reed et al. 2017; Andika et al. 2023b), or extreme deconvolution (e.g., Nanni et al. 2022).

Magnification of high-redshift quasars due to gravitational lensing by a massive galaxy on the same line of sight enables the study of quasars of intrinsically lower luminosity, and yields insights into their accretion disks (e.g., Chan et al. 2021) and host galaxies (e.g., Stacey et al. 2018). Observations of lensed quasars also place constraints

⋆ E-mail: ajnb3@ast.cam.ac.uk

on the dark matter profiles in the lensing galaxies (e.g., Gilman et al. 2020), and time-delay cosmography can be used to estimate the Hubble constant independently of the discrepant early- and late-Universe measurements (Refsdal 1964; Treu et al. 2022).

However, it has long been known that traditional photometric selection criteria exclude lensed high-redshift quasars, as flux from the galaxy can pollute the quasar colour to the extent that it escapes selection (e.g. Wyithe & Loeb 2002, Pacucci & Loeb 2019). Indeed, the first (and to date, only confirmed) strongly-lensed $z > 6$ quasar to be discovered, J0439+1634 at $z = 6.51$, had such a faint lensing galaxy that if the lens were only 0.5 mag brighter, the quasar would not have met the selection criteria applied (Fan et al. 2019). As a result, the fraction of high-redshift quasars observed to be lensed is significantly lower than expected. Indeed, the high lensing optical depth at $z > 6$ means that the lensed fraction is likely to be higher, with various studies suggesting it should be at least about 1 per cent (Yue et al. 2022a) or as high as 20 per cent (Pacucci & Loeb 2019). It is therefore likely that the selection criteria used to identify high-redshift quasars are excluding an entire population.

This paper is not the first attempt to select lensed quasars. Andika et al. (2023a) use an ensemble of state-of-the-art convolutional neural networks (CNNs) trained on mock images to identify 3080 candidate lensed quasars at $z > 1.5$ in the Hyper Suprime-Cam Subaru Strategic Survey (HSC-SSP; Aihara et al. 2022), pared down by visual inspection and astrometric information to 210, awaiting spectroscopic confirmation. At $z \gtrsim 6$, Andika et al. (2023b) use a combination of CNNs and SED fitting to identify 448 lensed and unlensed candidate high-redshift quasars in DES, reduced by visual inspection to 36, which also await confirmation. Yue et al. (2023) train a probabilistic random forest for their colour selection criteria, followed by morphological selection and a visual inspection phase, discovering an intermediately-lensed (magnified but not multiply-imaged) quasar as well as a quasar pair at $z \sim 5$.

The above are just three examples of the numerous and varied applications which machine learning (ML) is finding in the natural sciences. However, the reliance of many ML techniques on accurately labelled training data often makes such methods impractical. In this specific case, the absence of a large training set of real lensed $z \sim 6$ quasars is a key weakness of supervised methods. Unsupervised ML circumvents this by comparing instances to each other and grouping them according to similarity. In particular, contrastive learning (e.g. Chen et al. 2020) seeks a low-dimensional representation of the data by training a CNN to project innocuous transformations (rotations, reflections, etc.) of the same image closer together in a latent space, while aggressively separating projections of different sources.

This paper presents a new contrastive-learning-based methodology to find high-redshift quasars, lensed or not. We demonstrate the efficiency of this new method in isolating an 'island' of 11 sources, of which 10 are quasars. This includes three new high-redshift quasars confirmed spectroscopically, two of which may have been missed by previous searches due to their detection in the g band or their $i - z$ colour. One source spectrum exhibits flux below the Lyman limit, unambiguously indicating the presence of a foreground source which may be lensing the background quasar. This paper is structured as follows. In Section 2 we describe our candidate selection criteria, the contrastive learning approach, and SED fitting of our targets. Section 3 presents the observations confirming these targets as high-redshift quasars. We discuss implications of these observations in Section 4, and summarize our work in Section 5.

Magnitudes reported in this paper are in the AB system. Where required, we use a flat cosmology with $H_0 = 70 \, \text{km} \, \text{s}^{-1} \, \text{Mpc}^{-1}$ and $\Omega_{m0} = 0.3$.
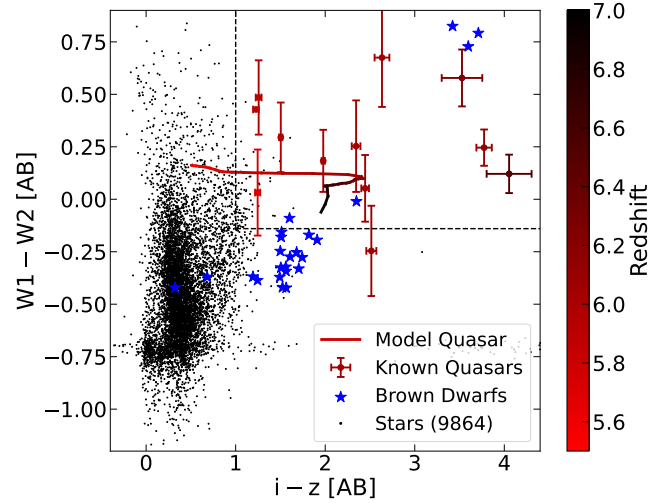
Figure 1. Photometric selection criteria. The black points are sources selected from three random DES tiles, most of which are likely stars. The blue stars are brown dwarfs from Kirkpatrick et al. (2011) and Best et al. (2015) cross-matched to DES. Several known high-redshift quasars are plotted, as well as a model quasar track generated using QSOGen (Temple et al. 2022); both are coloured according to redshift. Our liberal photometric selection criteria are shown by the dashed lines: $i - z > 1$ and $W1 - W2 > -0.14$ (both AB).

2 METHODS & TARGET SELECTION

2.1 Data preparation

We have used photometric and imaging data from the DES Data Release 2 (DR2), covering 5000 deg^2 of the southern celestial hemisphere in the broad $grizY$ bands (Abbott et al. 2021). The DES DR2 catalogue contains photometry and imagery for almost 700 million objects. Here we outline the criteria used to select our working sample; some of these criteria are also shown in Fig. 1. Our photometric selection criteria are less stringent than in previous studies in order to accommodate lensed quasars, where the lensing galaxy is expected to alter the quasar colour. All $grizY$ magnitudes quoted in this work use the DES DR2 aperture 4, equivalent to 1.95 arcsec diameter (Abbott et al. 2021).

We first select sources with $z < 21.0$ and $Y < 22.45$; we also impose magnitude errors of < 0.1 on each of these bands. We further select sources with $i - z > 1.0$ to preferentially select quasars and eliminate contaminants such as stars and galaxies, though this is a lower cutoff than in many previous searches (e.g. Reed et al. 2015, 2017; Wang et al. 2019; Bañados et al. 2016, 2023) to allow for photometric contamination from a lensing galaxy. Unlike most previous work, no requirement is placed on the g or r bands. We apply further standard cuts on the `IMAFLAGS_ISO` and `FLAGS` parameters to remove artefacts such as cosmic rays and saturation trails (Abbott et al. 2021). The DES DR2 SQL query combining these criteria can be found in Appendix A. These selection criteria result in 218 241 objects.

We further require a counterpart to each DES source in both VHS DR5 (McMahon et al. 2021) and AllWISE (Cutri et al. 2021) within 1 arcsec, to remove contaminants as well as artefacts that would be present in only one survey. After cross-matching to both surveys, 116 499 objects remain. We then use a colour cut of $W1 - W2 > -0.14 \, \text{mag}_{AB}$ ($W1 - W2 > 0.5 \, \text{mag}_{Vega}$), to remove a large number of cool dwarf stars, whose $grizY$ photometry is often similar to that of high-redshift quasars (e.g. Carnall et al. 2015). After further imposing
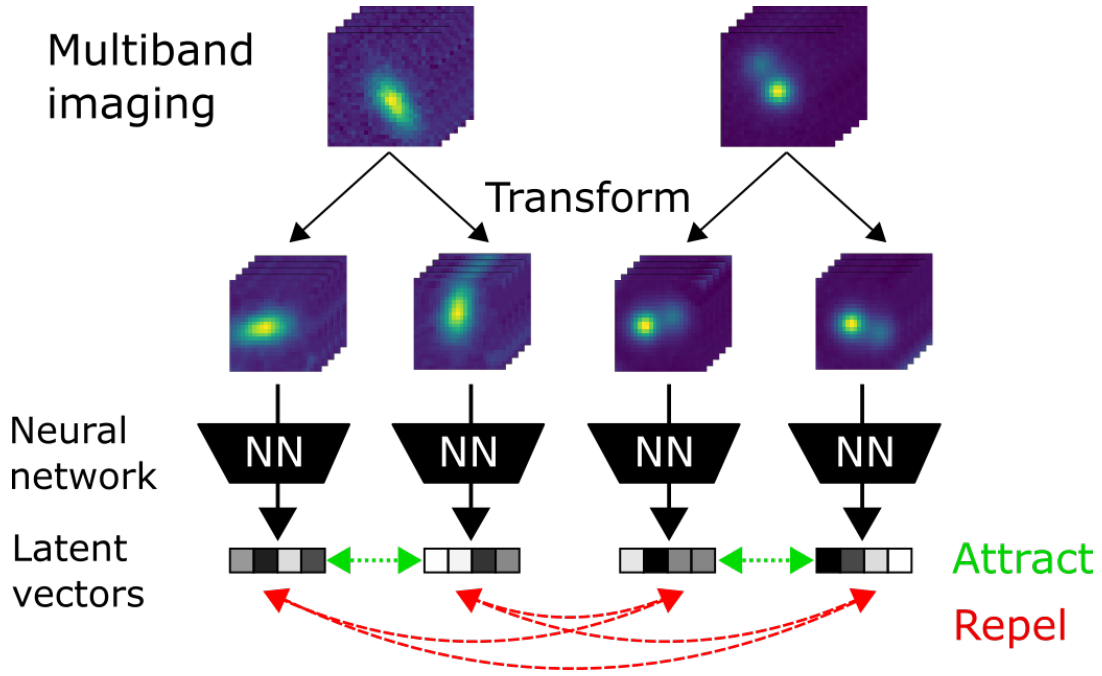
Figure 2. Schematic of contrastive learning. A batch of *N grizy* images are each randomly transformed twice, before being inputted into a convolutional neural network. The output of this network is a set of $2N$ vectors in a latent space. The contrastive loss function, computed using the locations of these vectors in the latent space, encourages proximity between vectors from the same source (the green dotted "attract" arrows), and discourages proximity between different sources (red dashed "repel" arrows). The training procedure is illustrated above for a batch size of 2; in practice a larger batch size is used (we use 128, see Table B1). The high-dimensional latent vectors output by the neural network are represented by the shaded sets of 4 boxes; in practice the latent dimension is larger (we use 128).

detection in *J*, *K*, *W1* and *W2* bands, with S/N > 2 in all bands, there are 7 438 objects remaining.

Despite the large number of remaining objects, no further colour cuts are made, to prevent lensed quasars from being excluded. Traditional photometric selection criteria are usually more restrictive, requiring higher thresholds for the $i - z$ colour, making extra cuts on the $z - Y$ and $Y - J$ colours, and requiring non-detections in the *g* and *r* bands (e.g., Bañados et al. 2014, 2016). Some also take steps to exclude extended sources at this stage (e.g., Reed et al. 2015), but we delegate morphology-based selection to our ML procedure (see Section 2.2). Stricter criteria naturally lead to a significantly smaller sample size ($\sim 10^2$ objects in the DES–VHS–WISE footprint, e.g. Reed et al. 2015, 2017) from which it is not impractical to simply inspect the sources individually or run SED-fitting procedures to select good quasar candidates. However, one of the aims of this work is to dispense with visual inspection altogether.

Furthermore, given the relatively low number density of high-redshift quasars (e.g. Schindler et al. 2023a) it is likely that our sample of 7 438 objects is dominated by contaminants, even taking into account a potentially large fraction of lensed quasars at $z \sim 6$ (e.g., Pacucci & Loeb 2019; Yue et al. 2022b). We suggest that it is difficult and perhaps impossible to remove these contaminants by a process based solely on photometry without also removing lensed high-redshift quasars: it is likely that they occupy the same regions of colour space. Indeed, Fig. 1 shows that the $i - z$ and $W1 - W2$ colours of high-redshift quasars are shared by brown dwarfs and other objects (such as dusty low-redshift galaxies).

Instead, we suggest that it is possible in principle to distinguish high-redshift quasars from contaminants using source imaging, another data product of DES (Flaugher et al. 2015). The large number of candidates motivates the use of an ML-based selection, partic-

ularly as the application of ML to image classification has rapidly improved in recent years (e.g. LeCun et al. 2015; Pak & Kim 2017).

2.2 Contrastive Learning

A requirement of supervised ML techniques is the availability of accurate ground-truth labels for the data: poorly-labelled data can severely impede image classification (Frenay & Verleysen 2014). Although several known high-redshift quasars are present in our sample, it is *a priori* unknown whether lensed quasars would appear substantially different in DES imaging. The only known lensed high-redshift quasar (J0439+1634) has been described in detail (Fan et al. 2019), but is not in the DES footprint. Additionally, the serendipity of its discovery and its high magnification ($\mu \sim 50$) suggest that its appearance could be unrepresentative of the lensed high-redshift quasar population: magnification distributions are typically taken to be $\propto \mu^{-3}$ (Schneider et al. 1992; Wyithe et al. 2011; Yue et al. 2022b). It would therefore be inappropriate to label objects as lensed high-redshift quasars based on subjective similarity to J0439+1634. In any case, it would be impractical to individually inspect and attempt to label each image in our large working sample.

Unsupervised ML techniques do not require ground-truth data labels, circumventing the above issues. Not only does this remove the necessity to label each image by hand, it avoids biases due to any prior assumptions about the appearance of lensed quasars in DES imaging.

The novelty in our candidate selection procedure is the use of contrastive learning, an unsupervised ML technique (Chen et al. 2020). Briefly, this method trains a neural network to group together similar images in a high-dimensional latent space. The loss function used to train the network encourages proximity in the latent space between

projections of transformations of the *same* images, and penalizes proximity of transformations of *different* images. Over the course of the self-supervised training, similar objects thus end up being clustered together in the latent space. A schematic of this technique is shown in Fig. 2 and described below; the training hyperparameters are given in Table B1 in Appendix B.

The contrastive learning procedure is as follows. A batch of N images is selected at random from the data; each "image" consists of 28×28 pixel cutouts in the 5 $grizY$ bands of DES (7.5×7.5 arcsec), normalized to the brightest pixel across all five bands; we use here the fact that the photometric calibrations of the DES images are uniform Burke et al. (2018). Two sequences of random transformations are then made of each image before projection with a CNN. These transformations are such that the classification of the image should not be affected: the image is cropped and may be flipped, rotated, and translated, but the relative brightness of the source in the different bands – that is, the colour – is not changed, as this can be a diagnostic feature. Similarly, the image is not sheared or zoomed in any way, as this could (for example) transform an image of a quasar into an image resembling an extended galaxy. This transformation stage ensures that the learning process uses only information relevant to source classification, such as colour, shape, extension; not e.g. orientation, centering.

These $2N$ transformed images are then each inputted into a CNN, whose architecture is inspired by those of Sarmiento et al. (2021) and Chen et al. (2020), and given in Table B2 in Appendix B. Contrastive learning CNNs consist of three modules: data augmentation, a base encoder, and a projection head. The data augmentation module transforms each image as discussed above: randomly cropping to 24×24 px; randomly flipping it horizontally and/or vertically; randomly translating it by 2 px; randomly rotating it. For a given object, the same set of transformations is applied to the images in each band. The base encoder consists of a sequence of three 2D convolutional layers (Fukushima 1979; LeCun et al. 2015) with exponential linear unit activation (Clevert et al. 2015); each layer is followed by a max-pooling layer with pool size 2 (Weng et al. 1992). The output of this module is reshaped to a 512-dimensional vector. Finally, the projection head consists of three densely-connected layers, outputting a vector z_i in a high-dimensional (in this case 128-dimensional) latent space.

The training process consists of altering the weights of the neural network so that similar images are projected to nearby vectors in the latent space, and different images are projected far away from each other. The contribution of each pair of representations to the loss function is

$$L_{i;j} = -\log \frac{\exp(\hat{z}_i \cdot \hat{z}_j / \tau)}{\sum_{k=1, k \neq i}^{2N} \exp(\hat{z}_i \cdot \hat{z}_k / \tau)} \tag{1}$$

where $\hat{z}_i = z_i / \|z_i\|_{\mathrm{L2}}$, the vector z_j is that which came from the same original DES object as the vector z_i, and τ is a 'temperature' hyperparameter affecting the clustering of the representations. The loss for a batch is then the average of the above contributions over all $2N$ output vectors. In minimizing the loss function over multiple batches (and hence transformations), the network thus learns to group together latent vectors from different transformations of the same source, while separating all others.

Following training, the projection head is discarded and the images are inputted directly to the base encoder; Chen et al. (2020) showed that removing the projection head improves performance by over 10 per cent. The output of the encoder is a 512-dimensional vector for each object. We then apply dimensionality reduction to this set of vectors using t-distributed Stochastic Neighbour Embedding (t-SNE; van der Maaten & Hinton 2008), to be able to visualize the latent space in two dimensions while approximately preserving the distances between the vectors.

2.3 Application to DES quasar candidates

The result of the unsupervised clustering and dimensionality reduction is shown in Fig. 3, separating the sample defined in Section 2.1 into several defined clusters. An 'island' towards the right of the frame contains 11 objects, including 7 known high-redshift quasars, which we term 'Quasar Island'. Follow-up spectroscopy of the other four objects (see Section 3) showed three to be newly discovered high-redshift quasars, and one to be a cool dwarf. We summarise the properties of these four objects in Tables 1 and 2. One of these sources (J0109–5424) was likely discarded by previous searches due to its detection in the g band; this detection indicates contamination by a foreground object that will be discussed in Section 4.2.

Determining the detailed nature of the other groups is beyond the scope of this work, but we briefly mention their main characteristics. By inspection of a few candidates, the largest 'continent' appears to contain isolated, compact sources: most of these objects are likely to be cool dwarfs; the lower left edge of this continent contains many L and T dwarf candidates from dal Ponte et al. (2023). The three other islands towards the bottom and right of the frame contain sources with extended morphologies (from visual inspection), indicating that they are galaxies. They also appear to be separated according to distinct $grizY$ colours. We note that three peripheral members of these other islands are known high-redshift quasars: the one just below the quasar island (J0422–1927, Bañados et al. 2016), and two in the lower right inset (J0203+0012, Venemans et al. 2007; J0008–0626, Bañados et al. 2016). This highlights the fact that this method does not identify all the quasars in the sample, so there may well be some unidentified quasars that have been missed by our method.

3 FOLLOW-UP CONFIRMATION AND ARCHIVAL SEARCH

3.1 An archival $z \sim 6$ quasar: J0122–4609

We cross-matched the four unconfirmed candidates with the ESO archive and Gemini-South archive. We found that J0122–4609 had unpublished EFOSC2/NTT spectroscopy (programme 098.A-0439(A), PI: McMahon). It was observed on the 22nd/23rd December 2016 for 1200 s using the 1.0 arcsec-wide slit with grism #16 and the OG530 blocking filter. The spectrum was reduced with the PYPEIT package (Prochaska et al. 2020), using the associated bias, flats and standard returned by the ESO archive. The final spectrum is flux-calibrated and re-scaled to match the DES z-band magnitude (see Table 2). The spectrum clearly shows the Lyman break and blue rest-frame UV colour of a high-redshift quasar (see Fig. 4, second panel). We fit the spectrum of J0122–4609 with the 'median' quasar template of Bañados et al. (2016) deriving a redshift of $z = 5.99$.

3.2 Discovery of two new quasars at $z \sim 6$, including a lensed candidate

Two further candidates on 'Quasar Island', J0603–3923 and J0109–5424, were observed respectively on the 6th and 11th December 2022 with the Gemini Multi-Object Spectrograph (GMOS) South (Programme ID GS-2022B-FT-208, PI: Farina). The two candidates were observed in long-slit mode for 4×300 s exposures using the
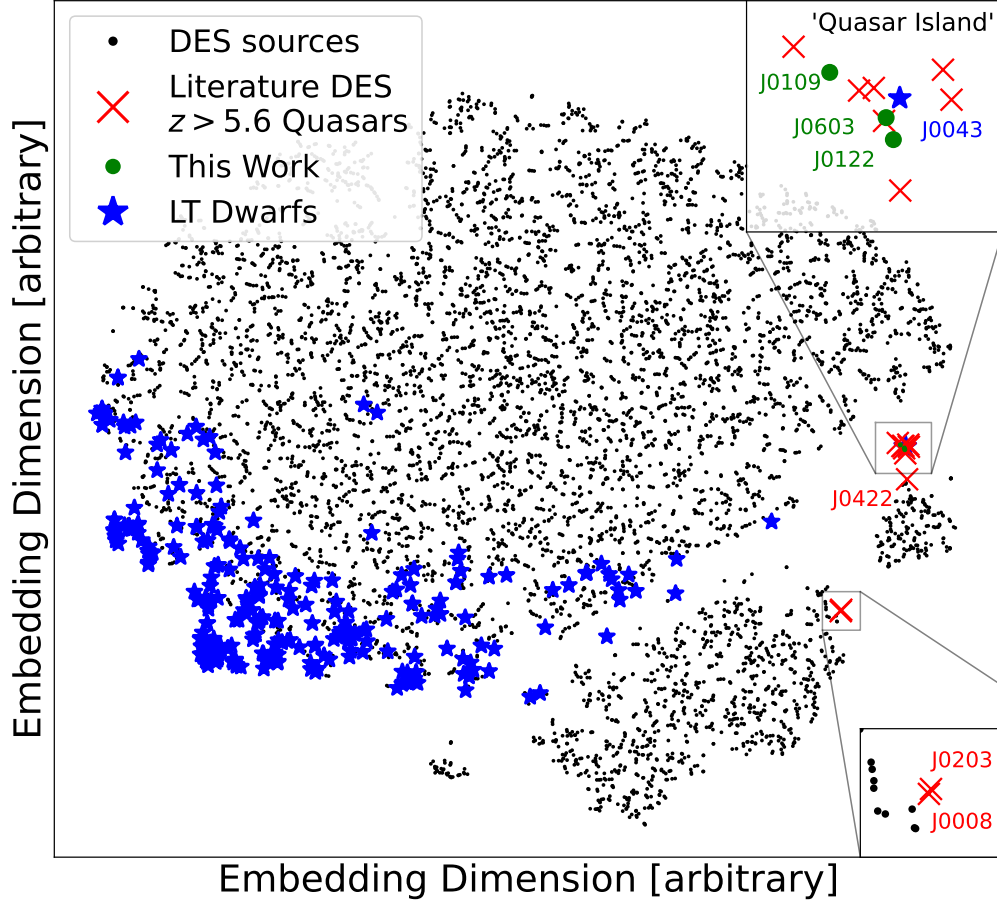
Figure 3. Dimensionality-reduced projection of DES images. The *t*-SNE procedure reduces the dimensionality of the set of vectors output by the encoder of the CNN while attempting to preserve the relative distances between each pair of points; as such the locations of the points would be equally valid if the plot were arbitrarily rescaled or rotated. There are several distinct groups of objects, including an isolated island towards the right of the frame, which we call 'Quasar Island'. This island contains 11 objects, of which seven are known high-redshift quasars (red crosses), three are newly-discovered quasars (green circles), and one is a brown dwarf (blue star). Three other literature high-redshift quasars are peripheral members of other islands: the one projected just below 'Quasar Island' and two in the lower right inset. The blue stars mostly at the lower left edge of the main continent are LT dwarf candidates from dal Ponte et al. (2023). The three other islands towards the bottom and right of the frame contain sources with extended morphologies, likely galaxies.

Table 1. Coordinates of the 11 objects clustered together on 'Quasar Island', and the three known high-redshift quasars not on this island. The first four are investigated by this work, with the neural network suggesting similarity between these and the seven other high-redshift quasars discovered by other authors. Redshifts are obtained spectroscopically, except for J0043–6028 which is a cool dwarf. The absolute magnitude at 1450 Å is computed using the best-fit quasar spectrum (see Fig. 5) and the Temple et al. (2022) template rest-frame UV slope $\alpha_\nu = -0.341$.

Object	R.A.	Dec.	Redshift	M_{1450}	Discovery
J0109–5424	$01^h09^m09\overset{s}{.}00$	$-54°24'16\overset{''}{.}29$	6.074	-26.52	This work
J0603–3923	$06^h03^m52\overset{s}{.}26$	$-39°23'35\overset{''}{.}78$	5.941	-25.83	This work
J0122–4609	$01^h22^m57\overset{s}{.}70$	$-46°09'14\overset{''}{.}20$	5.986	-26.51	This work
J0043–6028	$00^h43^m59\overset{s}{.}67$	$-60°28'44\overset{''}{.}85$	-	-	This work
ATLAS J029.9915–36.5658	$01^h59^m57\overset{s}{.}98$	$-36°33'56\overset{''}{.}67$	6.020	-27.0	Carnall et al. (2015)
VDESJ0454–4448	$04^h54^m01\overset{s}{.}79$	$-44°48'31\overset{''}{.}05$	6.100	-26.5	Reed et al. (2015)
PSO J056.7168–16.4769	$03^h46^m52\overset{s}{.}05$	$-16°28'36\overset{''}{.}98$	5.990	-26.72	Bañados et al. (2016)
VDESJ0408–5632	$04^h08^m19\overset{s}{.}24$	$-56°32'28\overset{''}{.}76$	6.035	-26.51	Reed et al. (2017)
VDESJ0224–4711	$02^h24^m26\overset{s}{.}55$	$-47°11'29\overset{''}{.}36$	6.526	-26.93	Reed et al. (2017)
VDESJ0143–5545	$01^h43^m10\overset{s}{.}25$	$-55°45'10\overset{''}{.}71$	6.250	-25.65	Reed et al. (2017)
PSO032.91882–17.07465	$02^h11^m40\overset{s}{.}53$	$-17°04'28\overset{''}{.}77$	6.050	-25.80	Bañados et al. (2023)
PSO J065.5041–19.4579	$04^h22^m01\overset{s}{.}00$	$-19°27'28\overset{''}{.}69$	6.125	-26.62	Bañados et al. (2016)
ULAS J0203+0012	$02^h03^m32\overset{s}{.}38$	$+00°12'29\overset{''}{.}06$	5.709	-26.2	Venemans et al. (2007)
PSO J002.1073–06.4345	$00^h08^m25\overset{s}{.}77$	$-06°26'04\overset{''}{.}42$	5.929	-26.0	Bañados et al. (2016)

Table 2. DR2 `MAG_APER_4` magnitudes of the 11 objects clustered together on 'Quasar Island', and the three known high-redshift quasars not on this island. All magnitudes are in the AB system. Upper limits are given at the 2σ level.

| Name	g	r	i	z	Y	J	K	W1	W2
J0109–5424	25.07 ± 0.15	23.09 ± 0.03	22.73 ± 0.04	20.18 ± 0.01	20.19 ± 0.03	19.53 ± 0.11	19.01 ± 0.15	19.12 ± 0.06	18.72 ± 0.08
J0603–3923	> 26.45	25.34 ± 0.19	22.43 ± 0.02	20.78 ± 0.01	20.85 ± 0.04	20.10 ± 0.16	20.08 ± 0.42	19.83 ± 0.09	19.64 ± 0.14
J0122–4609	> 26.45	24.28 ± 0.10	21.89 ± 0.02	20.09 ± 0.01	20.28 ± 0.03	19.73 ± 0.18	19.39 ± 0.24	19.41 ± 0.08	19.40 ± 0.14
J0043–6028	> 26.45	24.86 ± 0.16	22.15 ± 0.03	20.49 ± 0.01	20.52 ± 0.04	19.49 ± 0.06	20.02 ± 0.44	20.49 ± 0.18	19.84 ± 0.19
ATLAS J029.9915–36.5658	> 26.45	24.03 ± 0.07	21.86 ± 0.02	19.88 ± 0.01	20.03 ± 0.02	19.15 ± 0.10	19.17 ± 0.21	19.40 ± 0.08	19.21 ± 0.12
VDESJ0454–4448	> 26.45	25.51 ± 0.26	22.93 ± 0.04	20.49 ± 0.01	20.56 ± 0.03	20.32 ± 0.14	20.41 ± 0.46	19.68 ± 0.07	19.62 ± 0.14
PSO J056.7168–16.4769	> 26.45	24.27 ± 0.09	21.66 ± 0.02	20.15 ± 0.01	20.52 ± 0.04	19.89 ± 0.18	19.48 ± 0.30	19.50 ± 0.09	19.20 ± 0.14
VDESJ0408–5632	> 26.45	25.50 ± 0.33	22.69 ± 0.05	20.34 ± 0.01	20.38 ± 0.03	19.92 ± 0.13	19.48 ± 0.26	20.43 ± 0.12	20.18 ± 0.18
VDESJ0224–4711	> 26.45	> 26.15	24.56 ± 0.25	20.51 ± 0.01	20.37 ± 0.03	19.88 ± 0.14	19.00 ± 0.11	18.88 ± 0.05	18.76 ± 0.08
VDESJ0143–5545	> 26.45	> 26.15	24.38 ± 0.23	20.85 ± 0.02	21.28 ± 0.07	20.77 ± 0.27	19.77 ± 0.29	19.65 ± 0.09	19.07 ± 0.10
PSO032.91882–17.07465	> 26.45	> 26.15	23.61 ± 0.08	20.97 ± 0.02	21.08 ± 0.07	20.34 ± 0.16	19.40 ± 0.23	20.06 ± 0.16	19.39 ± 0.17
PSO J065.5041–19.4579	> 26.45	> 26.15	23.66 ± 0.09	19.88 ± 0.01	20.79 ± 0.04	19.83 ± 0.08	19.10 ± 0.15	18.67 ± 0.05	18.42 ± 0.07
ULAS J0203+0012	> 26.45	25.44 ± 0.29	22.16 ± 0.03	20.92 ± 0.01	21.11 ± 0.07	19.76 ± 0.14	19.23 ± 0.17	19.39 ± 0.09	19.36 ± 0.19
PSO J002.1073–06.4345	> 26.45	24.44 ± 0.16	21.98 ± 0.03	20.72 ± 0.02	21.05 ± 0.08	20.36 ± 0.11	19.83 ± 0.25	19.51 ± 0.11	19.02 ± 0.14
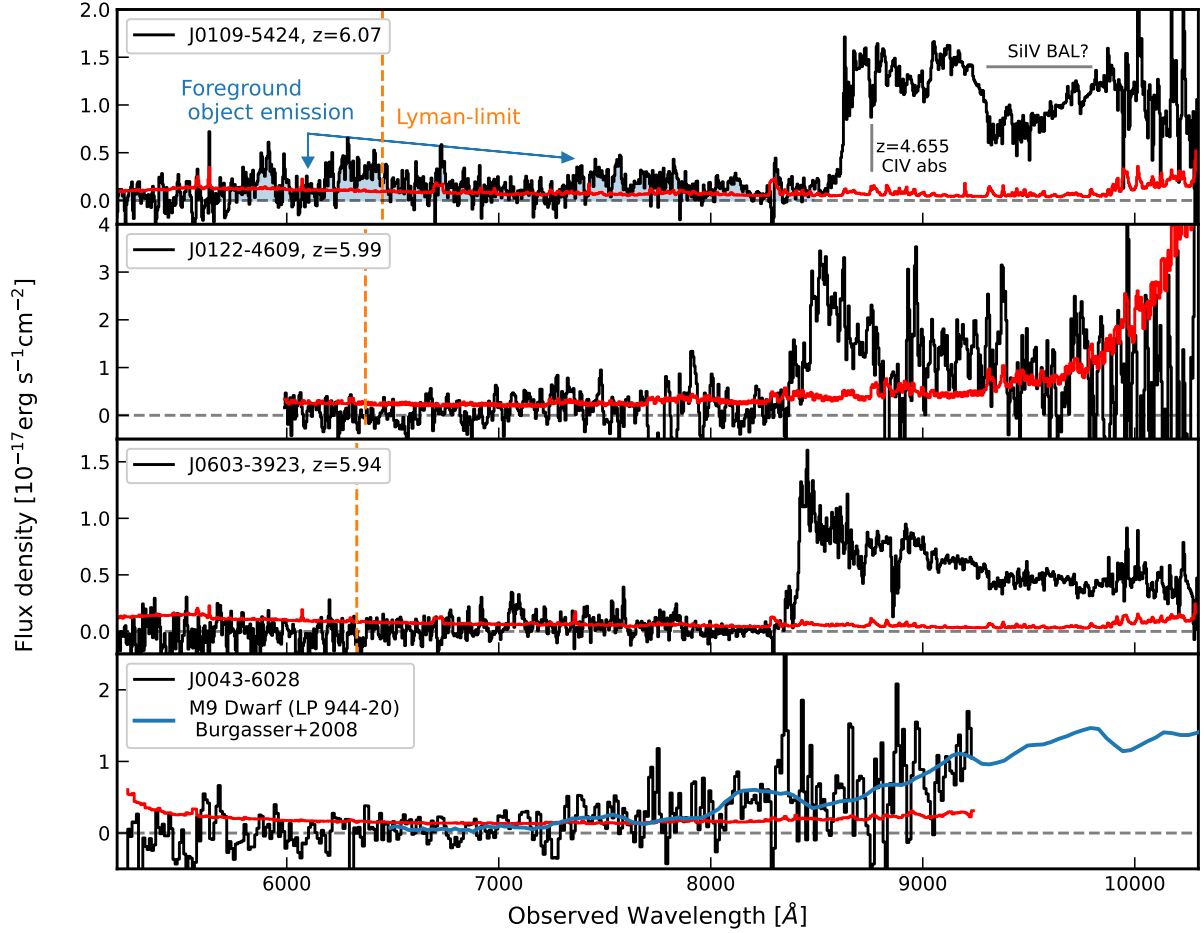
Figure 4. Confirmation spectra of the four unpublished objects of 'Quasar Island' studied in this work with EFOSC2/NTT and Gemini-South/GMOS. The spectra are all flux-calibrated and scaled to the DES z-band magnitude (see Table. 2). We show the extracted 1D error array in red. We confirm the discovery of three new quasars at $z \sim 6$, bringing the completeness of 'Quasar Island' to 91%. The remaining object is a cool dwarf whose spectrum is closely matching that of a M9 dwarf (Burgasser et al. 2008). For J0109–5424, we highlight in blue the presence of flux in the Lyman-α forest and below the Lyman limit, which we interpret as evidence for a foreground lensing galaxy (see Section 4.2). For the other two quasars, no flux is detected below the Lyman limit (orange dashed line).

1.5 arcsec slit width. We used the R400+G5325 grating in combination with the GG455_G0329 blocking filter to cover the wavelength range 5300 Å ≤ λ ≤ 10300 Å where the Lyman break of a $z \sim 6$ quasar would be expected. As with J0122–4609 above, the spectra were reduced using PYPEIT (Prochaska et al. 2020), and the final flux-

calibrated spectrum is re-scaled to match the DES z-band magnitude (see Table. 2).

We show the optical GMOS spectra of the two sources in Fig. 4, revealing a clear Lyman-α break typical of $z \sim 6$ quasars. J0109–5424 shows an excess of flux in the Lyman-α forest and below the Lyman Continuum limit consistent with the g and r band detection

which we interpret as evidence for a foreground lensing galaxy (see further Section 4.2). We fit J0603–3923 with the 'median' quasar template of Bañados et al. (2016) and J0109–5424 with the 'weak' template, deriving best-fitting redshifts of $z = 5.94$ and 6.07, respectively. The error of these templates compared to the Mg II or [C II] redshifts is typically $\sigma(\Delta z < 0.03)$ (Bañados et al. 2023). Both quasars are detected in the r band, and J0109–5424 also in the g band, which may have led them to be discarded as contaminants in traditional colour-cut searches. However, the GMOS spectra confirm their quasar nature, validating the efficiency of our contrastive network to find lensed quasar candidates.

3.3 The cool dwarf J0043–6028

The final object in 'Quasar Island' is J0043–6028, which also had archival unpublished EFOSC2/NTT spectroscopy (programme 0100.A-0346(A), PI: Reed), taken over 1800 s using grating #16 on the night of the 13/14 November. The achieved depth of these observations was too low to conclude on the nature of this object. Therefore, we re-observed this object using EFOSC2/NTT on the night of the 30 November / 1st December 2023 (programme 112.25VZ.001, PI: Belladitta) for 7200 s using grating #5. The data were reduced using PYPEIT (Prochaska et al. 2020) using the associated bias, flats and standard GD108 (observed during the night with the same grating).

We show the extracted 1D spectrum at the bottom of Fig. 4. The 1D spectrum is consistent that of a M9 dwarf star (Burgasser et al. 2008), consistent with the DES brown dwarf census using DES Y3, VHS and WISE photometry (Carnero Rosell et al. 2019).

4 DISCUSSION

4.1 Evaluation of contrastive learning methodology

Of the 11 objects selected by the contrastive network, 10 are high-redshift quasars and one is a cool dwarf star, giving an efficiency (true positive rate) of 91 per cent. At least three high-redshift quasars were not selected, giving a miss rate (false negative rate) of at least 23 per cent. Our methodology therefore has not identified all of the quasars in the sample; it is possible that further high-redshift quasars in the DES–VHS–WISE footprint remain undiscovered. Other machine learning methods may be able to provide a more complete selection.

We have managed to dispense with an extensive visual inspection stage in our candidate selection. Although not excessively time-consuming for $\sim 10^2$ objects, this would have been tedious with our much larger working sample of 7 438, obtained with the broader selection criteria we require to include lensed quasars. Aside from requiring significant expert time, visual inspection procedures are unlikely to be entirely systematic and objective. The lack of this stage in our methodology therefore aids its reproducibility.

4.2 J0109–5424 as a candidate lensed high-redshift BAL quasar

Whilst the Lyman-α line and redwards continuum of J0603–3923 is typical of high-redshift quasars, J0109–5424 presents strong absorption features and weak Lyman-α emission (see Fig. 4). The strong absorption at $\sim 9100 - 9800$ Å is consistent with Si IV broad absorption line (BAL) outflow. The shape of the BAL is strongest at high velocities (up to ~ 18000 km s^{-1}), which is unusual in BALs at high- and low-redshifts (e.g. Weymann et al. 1991; Trump et al. 2006; Bischetti et al. 2022). A near-infrared spectrum confirming the presence of the associated C IV BAL and/or contamination of the presumed

Si IV BAL other absorption features, as well as characterising the shape of the rest-frame UV continuum up to and beyond the C IV line is necessary to conclude. We also identify a C IV absorption system at $z = 4.655$, although a higher resolution spectrum would be necessary to confirm it.

Most interestingly, the spectrum of J0109–5424 presents significant flux throughout the Lyman-α forest and even below the Lyman limit. This is an unambiguous signature of a foreground, potentially lensing, galaxy contaminating the photometry of the quasar, making J0109–5424 the first lensed BAL quasar candidate at $z > 5.5$. Further, the SExtractor *class_star* probability (0: extended source, 1: point-source) for the g band is 0.32, suggesting that the contaminating object is likely a galaxy (for the $rizY$ bands the probabilities are respectively 0.82, 0.96, 0.98, 0.84, consistent with the point-source nature of the quasar).

SED modelling also suggests J0109–5424 to be lensed, unlike J0603–3923 or J0122–4609. We fit the photometry of these three quasars to templates of galaxies, quasars, and lensed quasars, using a custom code employing the EMCEE Python package (Foreman-Mackey et al. 2013). The galaxy models were constant-star-formation-rate models from the BAGPIPES package (Carnall et al. 2018), parametrized by stellar mass, start and end times of star formation, metallicity, redshift, and dust extinction. The quasar templates were generated using QSOGEN (Temple et al. 2021, 2022), parametrized using a flux-free scaling factor and colour excess E_{B-V}, and the IGM absorption prescription from Becker & Bolton (2013)[1]. Finally, we create simple 'galaxy+quasar' composite templates by adding the galaxy templates to the quasar templates, in order to mimic to first order the expected spectrum of a lensed quasar system. The galaxy templates have six free parameters; the quasar templates two; the lensed quasar templates eight. To quantify goodness of fit for models with different numbers of fitting parameters, we use the Bayesian information criterion (BIC; Schwarz 1978), which penalizes models with more parameters to discourage overfitting. The best-fitting spectra for the three quasars against the three templates are shown in Fig. 5, along with the BICs for the each candidate and model. For J0109–5424, we find that a galaxy+quasar model is a significantly better model than a pure quasar model by ΔBIC > 150. The left panel of Fig. 5 shows that the inclusion of galaxy flux in the modelling adds sufficient flexibility to fit the flux measured in the g band. For the other two objects, the best-fitting galaxy+quasar model is essentially identical to the best-fitting quasar model: the galaxies in these models contribute very little flux across all bands. These quasars have a higher BIC for galaxy+quasar models than pure quasar models; we therefore consider these objects to be unlensed quasars.

Another argument for the lensed nature of J0109–5424 is its extremely short proximity zone. Using the standard definition of the ionised near-zone (Eilers et al. 2021), we measure $r_{NZ} = 0.41 \pm 0.12(1.15)$ Mpc, where the first error is due to the resolution of the GMOS spectra and the second stems from the uncertainty of the quasar redshift. The expected proximity zones for J0109–5424 ($M_{1450} = -26.52$) is $r_{NZ} \simeq 4$ Mpc, although considerable scatter exists in the $z = 6$ quasar population (Eilers et al. 2017,

---

[1] We note that QSOGEN uses the IGM absorption prescription of Becker & Bolton (2013) which results bluer $i - z$ colours for $z > 5.8$ quasars (e.g. due to less Lyman-α forest absorption) than observed (see Appendix B of Schindler et al. 2023b, for more details). In our specific case, this makes the modelling of a combined foreground galaxy and background quasar spectrum more conservative as the inclusion of a lensing galaxy would also make the $i - z$ colour bluer.
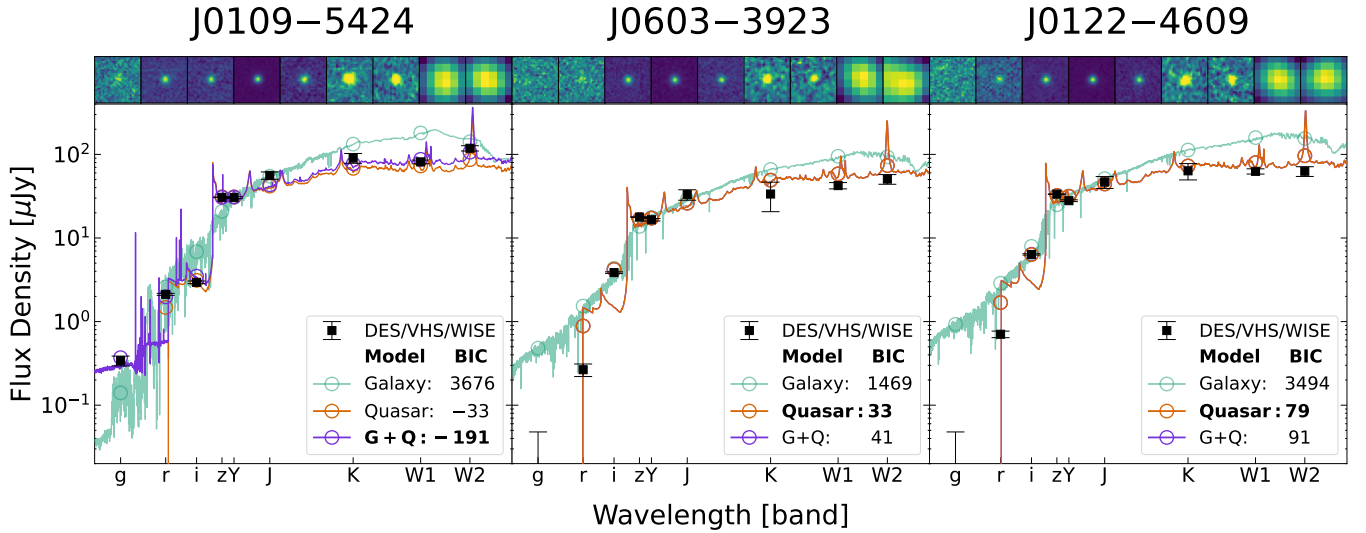
Figure 5. DES+VHS+WISE imaging and best-fitting spectra of the candidates' photometry for galaxy, quasar and lensed quasar templates based on BAGPIPES galaxy models (Carnall et al. 2018) and QSOGEN quasar models (Temple et al. 2021, 2022) as described in the text. The images are each 7.5 " across. Integrated fluxes for *grizY*, *J*, *K*, *W1* and *W2* bands are indicated in the coloured circles. The observed fluxes measured by the three surveys are indicated by the black squares with error bars. The BICs for the model fits are given in the legend. For J0109–5424, a galaxy+quasar template provides a significantly better fit than a pure galaxy or quasar template, and the source is discernable in the *g* band imaging (top row, far left). For J0603–3923 and J0122–4609, the best-fitting galaxy+quasar models are almost identical to the best-fitting quasar models, and there is no clear source in the *g* band.

2021; Satyavolu et al. 2023). However, the near-zone of J0109–5424 is much smaller than expected, and is more common for quasars with $-23 \lesssim M_{1450} \lesssim -24$ (depending on the lifetime assumed, e.g. Satyavolu et al. 2023)). This strongly argues for this object to be gravitationally lensed and magnified (see e.g. Davies et al. 2020). We can roughly estimate the total magnification to $10 < \mu < 25$, but a more precise redshift (for example using FIR lines targeted with ALMA or rest-frame optical lines with JWST) are needed with a higher-resolution optical spectrum to improve the proximity zone measurement. Space-based high-resolution imaging is however necessary to definitely confirm the lensed nature of J0109-5424 by detecting multiple images.

In summary, multiple arguments point to a lensed nature for J0109–5424: the detection of flux below the Lyman limit, the SED modelling, and the near-zone size.

4.3 Implications for the lensing fraction at $z \sim 6$

Confirming the lensed nature of J0109–5424 would double the observed lensed fraction of $z \sim 6$ quasars. The DES–VHS–WISE detected sample we have used in this work contains 13 $z \sim 6$ quasars (see Section 2.2), of which one is a lensed quasar candidate. This would imply an observed lensed fraction of $7.7^{+10.8}_{-4.7}$ per cent. This is consistent with both Pacucci & Loeb (2019) who predict a $z \sim 6$ lensed fraction between 6 and 20 per cent, and Yue et al. (2022b) who give a revised estimate using more recent lens parameters at 1–6 per cent depending on the bright-end slope of the $z \sim 6$ quasar luminosity function and the depth of the survey considered. Compared to the several hundred high-redshift quasars discovered to date, the lensed fraction we obtain predicts ~ 10 quasars at $z \sim 6$ to be lensed, suggesting that several are yet to be discovered in archival ground-based imaging survey data.

5 CONCLUSIONS

We have developed a novel methodology based on unsupervised machine learning to find high-redshift quasars that have been missed by previous surveys, including gravitationally-lensed objects. This new method, using relaxed colour cuts and DES imaging data, is completely automated and does not require any visual inspection. Our method isolated 11 objects in the DES–VHS–WISE footprint, 10 of which are high-redshift quasars, implying an efficiency of 91%. We note however that the selection is not complete: three known high-redshift quasars are located outside of 'Quasar Island'. Of these, we report the discovery of three new $z \sim 6$ quasars. One object, J0109–5424 at $z = 6.07$, is likely strongly lensed: it is detected in the *g* band, its spectrum shows significant flux in the Lyman-α forest and even below the Lyman limit, and the ionised near-zone is too small considering the apparent luminosity of the quasar. Detection in the *g* band had likely caused this quasar to be missed by previous searches. Follow-up high-resolution imaging is now needed to confirm the lensed nature of J0109–5424, which would make it the second lensed $z > 6$ quasar to be discovered.

The discovery of a lensed $z > 6$ quasar in the DES footprint would suggest that of order 10 similar quasars have been missed in previous searches due to the photometric contamination induced by lensing galaxies. The magnification effect induced by gravitational lensing brings ever fainter quasars into view, and this work suggests that lensed high-redshift quasars are not as uncommon as the paucity of the known population would suggest. As such it is likely that lensed quasars can be observed deeper into the epoch of reionisation, opening up new avenues with which to probe the high-redshift Universe.

erenced in the text, we also acknowledge the use of NumPy (Harris et al. 2020), Matplotlib Hunter (2007), PyVO (Graham et al. 2014), pandas (pandas development team 2020; Wes McKinney 2010), and Astropy (Astropy Collaboration et al. 2013, 2018, 2022).

DATA AVAILABILITY

The DES photometric and imaging data used in this work are publicly available at `https://www.darkenergysurvey.org/the-des-project/data-access/`. The contrastive network architecture and trained weights will be made available upon acceptance of the manuscript at `https://github.com/xbyrne/glq_mpia`. The NTT/EFOSC2 raw data of J0122–4609 and J0043–6028 presented in this work are available in the ESO archive (Proposal IDs: 098.A-0439(A), 112.25VZ.001). The Gemini GMOS spectra raw data are available in the Gemini Observatory Archive (Programme ID GS-2022B-FT-208).

REFERENCES

Abadi M., et al., 2015, TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems, https://www.tensorflow.org/
Abbott T. M. C., et al., 2021, ApJS, 255, 20
Aihara H., et al., 2022, PASJ, 74, 247
Andika I. T., et al., 2023a, A&A, 678, A103
Andika I. T., et al., 2023b, ApJ, 943, 150
Astropy Collaboration et al., 2013, A&A, 558, A33
Astropy Collaboration et al., 2018, AJ, 156, 123
Astropy Collaboration et al., 2022, ApJ, 935, 167
Bañados E., et al., 2014, AJ, 148, 14
Bañados E., et al., 2015, ApJ, 804, 118
Bañados E., et al., 2016, ApJS, 227, 11
Bañados E., et al., 2023, ApJS, 265, 29
Becker G. D., Bolton J. S., 2013, MNRAS, 436, 1023
Best W. M. J., et al., 2015, ApJ, 814, 118
Bischetti M., et al., 2022, Nature, 605, 244
Boch T., Pineau F., Derriere S., 2012, in Ballester P., Egret D., Lorente N. P. F., eds, Astronomical Society of the Pacific Conference Series Vol. 461, Astronomical Data Analysis Software and Systems XXI. p. 291
Bosman S. E. I., Davies F. B., Bañados E., 2023, Research Notes of the American Astronomical Society, 7, 72
Burgasser A. J., Liu M. C., Ireland M. J., Cruz K. L., Dupuy T. J., 2008, ApJ, 681, 579
Burke D. L., et al., 2018, AJ, 155, 41
Carnall A. C., et al., 2015, MNRAS, 451, L16
Carnall A. C., McLure R. J., Dunlop J. S., Davé R., 2018, MNRAS, 480, 4379
Carnero Rosell A., et al., 2019, MNRAS, 489, 5301
Chambers K. C., et al., 2016, arXiv e-prints, p. arXiv:1612.05560
Chan J. H. H., Rojas K., Millon M., Courbin F., Bonvin V., Jauffret G., 2021, A&A, 647, A115
Chen T., Kornblith S., Norouzi M., Hinton G., 2020, arXiv e-prints, p. arXiv:2002.05709
Clevert D.-A., Unterthiner T., Hochreiter S., 2015, arXiv e-prints, p. arXiv:1511.07289
Cutri R. M., et al., 2021, VizieR Online Data Catalog, p. II/328
Davies F. B., Wang F., Eilers A.-C., Hennawi J. F., 2020, ApJ, 904, L32
Eilers A.-C., Davies F. B., Hennawi J. F., Prochaska J. X., Lukić Z., Mazzucchelli C., 2017, ApJ, 840, 24
Eilers A.-C., Hennawi J. F., Davies F. B., Simcoe R. A., 2021, ApJ, 917, 38
Fan X., et al., 2001, AJ, 122, 2833
Fan X., et al., 2019, ApJ, 870, L11
Fan X., Bañados E., Simcoe R. A., 2023, ARA&A, 61, 373
Flaugher B., et al., 2015, AJ, 150, 150
Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, PASP, 125, 306
Frenay B., Verleysen M., 2014, IEEE Transactions on Neural Networks and Learning Systems, 25, 845
Fukushima K., 1979, IEICE Technical Report, A, 62, 658
Gilman D., Birrer S., Nierenberg A., Treu T., Du X., Benson A., 2020, MNRAS, 491, 6077
Graham M., Plante R., Tody D., Fitzpatrick M., 2014, PyVO: Python access to the Virtual Observatory, Astrophysics Source Code Library, record ascl:1402.004 (ascl:1402.004)
Harris C. R., et al., 2020, Nature, 585, 357
Hunter J. D., 2007, Computing in Science & Engineering, 9, 90
Inayoshi K., Visbal E., Haiman Z., 2020, ARA&A, 58, 27
Jiang L., et al., 2016, ApJ, 833, 222
Kirkpatrick J. D., et al., 2011, ApJS, 197, 19
LeCun Y., Bengio Y., Hinton G., 2015, Nature, 521, 436
McMahon R. G., Banerji M., Gonzalez E., Koposov S. E., Bejar V. J., Lodieu N., Rebolo R., VHS Collaboration 2013, The Messenger, 154, 35
McMahon R. G., Banerji M., Gonzalez E., Koposov S. E., Bejar V. J., Lodieu N., Rebolo R., VHS Collaboration 2021, VizieR Online Data Catalog, p. II/367
Nanni R., Hennawi J. F., Wang F., Yang J., Schindler J.-T., Fan X., 2022, MNRAS, 515, 3224
Pacucci F., Loeb A., 2019, ApJ, 870, L12
Pak M., Kim S., 2017, in 2017 4th International Conference on Computer Applications and Information Processing Technology (CAIPT). pp 1–3, doi:10.1109/CAIPT.2017.8320684
Prochaska J. X., et al., 2020, Journal of Open Source Software, 5, 2308
Reed S. L., et al., 2015, MNRAS, 454, 3952
Reed S. L., et al., 2017, MNRAS, 468, 4702
Refsdal S., 1964, MNRAS, 128, 307
Richards G. T., et al., 2002, AJ, 123, 2945
Sarmiento R., Huertas-Company M., Knapen J. H., Sánchez S. F., Domínguez Sánchez H., Drory N., Falcón-Barroso J., 2021, ApJ, 921, 177

Satyavolu S., Kulkarni G., Keating L. C., Haehnelt M. G., 2023, MNRAS, 521, 3108

Schindler J.-T., et al., 2023a, ApJ, 943, 67

Schindler J.-T., et al., 2023b, ApJ, 943, 67

Schneider P., Ehlers J., Falco E. E., 1992, Gravitational Lenses. Springer Berlin Heidelberg, Berlin, Heidelberg, doi:10.1007/978-3-662-03758-4_1, https://doi.org/10.1007/978-3-662-03758-4_1

Schwarz G., 1978, Annals of Statistics, 6, 461

Stacey H. R., et al., 2018, MNRAS, 476, 5075

Temple M. J., Hewett P. C., Banerji M., 2021, MNRAS, 508, 737

Temple M. J., Hewett P. C., Banerji M., 2022, QSOGEN: Model quasar SEDs, Astrophysics Source Code Library, record ascl:2205.003 (ascl:2205.003)

Treu T., Suyu S. H., Marshall P. J., 2022, A&ARv, 30, 8

Trump J. R., et al., 2006, ApJS, 165, 1

Venemans B. P., McMahon R. G., Warren S. J., Gonzalez-Solares E. A., Hewett P. C., Mortlock D. J., Dye S., Sharp R. G., 2007, MNRAS, 376, L76

Venemans B. P., et al., 2013, ApJ, 779, 24

Volonteri M., Habouzit M., Colpi M., 2021, Nature Reviews Physics, 3, 732

Wang F., et al., 2019, ApJ, 884, 30

Weng J., Ahuja N., Huang T. S., 1992, in International Joint Conference on Neural Networks (IJCNN). pp 576–581

Wes McKinney 2010, in Stéfan van der Walt Jarrod Millman eds, Proceedings of the 9th Python in Science Conference. pp 56 – 61, doi:10.25080/Majora-92bf1922-00a

Weymann R. J., Morris S. L., Foltz C. B., Hewett P. C., 1991, ApJ, 373, 23

Wright E. L., et al., 2010, AJ, 140, 1868

Wyithe J. S. B., Loeb A., 2002, ApJ, 577, 57

Wyithe J. S. B., Oh S. P., Pindor B., 2011, MNRAS, 414, 209

Yang D.-M., et al., 2022, arXiv e-prints, p. arXiv:2211.16996

York D. G., et al., 2000, AJ, 120, 1579

Yue M., Fan X., Yang J., Wang F., 2022a, ApJ, 925, 169

Yue M., Fan X., Yang J., Wang F., 2022b, ApJ, 925, 169

Yue M., Fan X., Yang J., Wang F., 2023, AJ, 165, 191

dal Ponte M., et al., 2023, MNRAS, 522, 1951

pandas development team T., 2020, pandas-dev/pandas: Pandas, doi:10.5281/zenodo.3509134, https://doi.org/10.5281/zenodo.3509134

van der Maaten L., Hinton G., 2008, Journal of Machine Learning Research, 9, 2579

APPENDIX A: SQL QUERY

Here we present the SQL code used to query the DES database.

```
SELECT
    COADD_OBJECT_ID, RA, DEC,
    MAG_APER_4_G, MAG_APER_4_R, MAG_APER_4_I,
    MAG_APER_4_Z, MAG_APER_4_Y,
    MAGERR_APER_4_G, MAGERR_APER_4_R,
    MAGERR_APER_4_I, MAGERR_APER_4_Z,
    MAGERR_APER_4_Y
FROM DR2_MAGNITUDE WHERE
    MAG_APER_4_Z < 21.0 AND
    MAGERR_APER_4_Z < 0.1 AND
    MAG_APER_4_Y < 22.45 AND
    MAGERR_APER_4_Y < 0.1 AND
    MAG_APER_4_I - MAG_APER_4_Z > 1.0 AND
    FLAGS_G < 4 AND FLAGS_R < 4 AND FLAGS_I < 4 AND
    FLAGS_Z < 4 AND FLAGS_Y < 4 AND
    IMAFLAGS_ISO_G = 0 AND IMAFLAGS_ISO_R = 0 AND
    IMAFLAGS_ISO_I = 0 AND IMAFLAGS_ISO_Z = 0 AND
    IMAFLAGS_ISO_Y = 0
```

Table B1. Training hyperparameters used for contrastive learning.

Hyperparameter	Value
Batch size (N)	128
Temperature (τ)	0.1
Optimizer	SGD
Learning Rate	3×10^{-4}
Momentum	1×10^{-5}
Patience	5
Minimum Delta	0.01

Table B2. CNN architecture. Layer names are those defined by TensorFlow (Abadi et al. 2015). All Conv2D and Dense layers have exponential linear unit activation (Clevert et al. 2015).

Module	Layer name	Hyperparameters	
Data Augmentation	RandomCrop	24×24px	
	RandomFlip		
	RandomTranslation	± 2px	
	RandomRotation	$\pm 180°$	
Base Encoder	Conv2D	128 filters	kernel 5
	MaxPooling2D		pool 2
	Conv2D	256	3
	MaxPooling2D		2
	Conv2D	512	3
	MaxPooling2D		2
	Reshape	512	
Projection Head	Dense	256 neurons	
	Dense	128	
	Dense	128	

APPENDIX B: CONTRASTIVE LEARNING TRAINING & ARCHITECTURE

The hyperparameters used in training the contrastive learning model are given in Table B1. The architecture of the model is given in Table B2.

The network was trained using the cloud-based GPUs offered by Google Colaboratory[2]. Early stopping was implemented with a patience of 5 and a minimum delta of 0.01, such that training is halted when the loss function fails to decrease by more than 0.01 after 5 epochs; this occurred after 88 epochs. Each epoch took approximately 8 s.

This paper has been typeset from a TeX/LaTeX file prepared by the author.

[2] https://colab.research.google.com