

The Mini-SiTian Array: the Mini-SiTian Real-time Image Processing Pipeline (STRIP) Post-print

Authors: Hongrui Gu, Yang Huang, Yongkang Sun, Kai Xiao, Zhirui Li, Beichuan Wang, Zhou Fan, Chuanjie Zheng, Henggeng Han, Hu Zou, Wenxiong Li, Hong Wu and Jifeng Liu

Date: 2025-05-09T08:40:03+00:00

Abstract

This paper provides a comprehensive introduction to the mini-SiTian Real-time Image Processing pipeline (STRIP) and evaluates its operational performance. The STRIP pipeline is specifically designed for real-time alert triggering and light curve generation for transient sources. By applying the STRIP pipeline to both simulated and real observational data of the Mini-SiTian survey, it successfully identified various types of variable sources, including stellar flares, supernovae, variable stars, and asteroids, while meeting requirements of reduction speed within 5 minutes. For the real observational data set, the pipeline detected one flare event, 127 variable stars, and 14 asteroids from three monitored sky regions. Additionally, two data sets were generated: one, a real-bogus training data set comprising 218,818 training samples, and the other, a variable star light curve data set with 421 instances. These data sets will be used to train machine learning algorithms, which are planned for future integration into STRIP.

Full Text

Preamble

Research in Astronomy and Astrophysics, 25:044007 (13pp), 2025 April © 2025. National Astronomical Observatories, CAS and IOP Publishing Ltd. All rights, including for text and data mining, AI training, and similar technologies, are reserved. Printed in China <https://doi.org/10.1088/1674-4527/adc790> CSTR: 32081.14.RAA.adc790

The Mini-SiTian Array: the Mini-SiTian Real-time Image Processing Pipeline (STRIP)

Kai Xiao^{1,2}, Beichuan Wang^{1,2}, Yongkang Sun^{1,2}, Yang Huang^{2,1,3}, Zhou Fan^{1,2,3}, Hongrui Gu^{1,2}, Zhirui Li^{1,2}, Chuanjie Zheng^{1,2}, Henggeng Han^{1,2}, Hu Zou^{1,2}, Wenxiong Li^{1,2}, Hong Wu^{1,2}, and Jifeng Liu^{1,2}

¹ National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100101, China; zfan@nao.cas.cn ² School of Astronomy and Space Science, University of Chinese Academy of Sciences, Beijing 100049, China; huangyang@ucas.ac.cn

Received 2024 November 11; revised 2025 March 4; accepted 2025 March 17; published 2025 May 6

Abstract

This paper provides a comprehensive introduction to the mini-SiTian Real-time Image Processing pipeline (STRIP) and evaluates its operational performance. The STRIP pipeline is specifically designed for real-time alert triggering and light curve generation for transient sources. By applying the STRIP pipeline to both simulated and real observational data of the Mini-SiTian survey, it successfully identified various types of variable sources, including stellar flares, supernovae, variable stars, and asteroids, while meeting requirements of reduction speed within 5 minutes. For the real observational data set, the pipeline detected one flare event, 127 variable stars, and 14 asteroids from three monitored sky regions. Additionally, two data sets were generated: one, a real-bogus training data set comprising 218,818 training samples, and the other, a variable star light curve data set with 421 instances. These data sets will be used to train machine learning algorithms, which are planned for future integration into STRIP.

Key words: surveys – techniques: photometric – stars: variables: general – techniques: image processing

1. Introduction

Transient sources are often associated with extreme physical processes. Investigating their early evolution can validate or expand our astrophysical theories under extreme astronomical phenomena, such as solar flares (Allred et al. 2015), kilonovae, and supernovae (Li et al. 2024). To study these transient events, time-domain surveys like the Lick Observatory Supernova Search (Li et al. 2000; Filippenko et al. 2001) began observing the sky toward the end of the last century. Subsequently, the emergence of projects such as the Palomar Transient Factory (Law et al. 2009; the predecessor to the currently operational Zwicky Transient Facility, ZTF, Bellm et al. 2019), Catalina Real-Time Transient Survey (Drake et al. 2009; Djorgovski et al. 2011), the Asteroid Terrestrial-impact Last Alert System (ATLAS; Tonry 2011; Tonry et al. 2018), the All-Sky Automated Survey for Supernovae (Shappee et al. 2014; significantly Kochanek et al. 2017), and other projects broadened human understanding of various types of transient

sources.

However, single-band time-domain surveys provide limited information. Since transient events are often non-repeating and short-lived, missing the early burst means that crucial temperature and spectral line features cannot be adequately captured using only single-band data. To address simultaneous multicolor observation limitations, numerous programs and high-cadence time-domain surveys have been developed to obtain more comprehensive information on transients. Examples include the Multi-channel Photometric Survey Telescope (Mephisto, Liu 2019), the Tsinghua University-M. Huateng Telescopes for Survey (Zhang et al. 2020), and the Ground Wide Angle Cameras Array (Han et al. 2021), which achieve minute-level sampling. However, existing wide-field, high-cadence telescope arrays generally have shallow detection limits, while large-aperture multicolor telescopes often lack sufficient sampling frequency. To overcome these challenges, the SiTian project (Liu et al. 2021) was initiated. Consisting of dozens of 1 m-aperture telescopes, each with a 25 square-degree field of view, SiTian is designed to capture simultaneous three-color images at a depth of magnitude 21 every 30 minutes, significantly enhancing our ability to study rapid and faint transient sources.

Nevertheless, time-domain surveys using only a few broadband filters provide far less information than spectroscopy. To fast capture the nature of transient sources, near-real-time processing of time-domain photometric images is essential, enabling a rapid response for multi-band follow-up observations, including optical spectroscopy, radio, and high-energy observations. In optical imaging, high-precision aperture photometry combined with catalog cross-matching can quickly and effectively identify variations of variable sources in non-crowded fields, thereby triggering follow-up spectroscopic observations or even observations in other wavelengths, such as radio and X-rays. However, many major transient sources, such as supernovae or tidal disruption events (hereafter TDEs), occur in crowded star fields or galactic disks, where aperture photometric measurements cannot be used. To address this issue, algorithms like ISIS (Alard & Lupton 1998; Alard 2000), Hotpants (Becker 2015), and SFFT (Hu et al. 2022) have been developed for image subtraction. Point-like features remaining in the post-subtraction image indicate variable sources, while the invariant, crowded stellar fields or galactic disks are removed, thus enabling the discovery and measurement of variable sources through aperture photometry.

The Mini-SiTian Array serves as the pathfinder for the SiTian project (Huang et al. 2025), utilizing the same SiTian project observational mode to function as an exploration and pre-research platform. To facilitate rapid responses to transients during observations, it is essential to have a dedicated transient response pipeline based on image subtraction techniques. This study, which is based on the Hotpants algorithm, has been tested on Mini-SiTian Field-02 (F02) and selected supernova fields. The tests have verified the pipeline's feasibility and robustness, successfully identifying a range of variable sources. The primary focus of this paper is to introduce the complete pipeline and demonstrate its effectiveness

through the results obtained from these tests.

The structure of this paper is as follows: Section 2 provides an overview of the workflow for the mini-SiTian Real-time Image Processing pipeline (STRIP). Section 3 evaluates the performance of the pipeline using both simulated and real observational data. Section 4 presents the data products generated from the test observations. Finally, Section 5 summarizes the key findings, discusses the accomplishments of the pipeline, and outlines the anticipated future directions for Mini-SiTian operations.

2. STRIP Pipeline

The Mini-SiTian telescopes are currently operational at the Xinglong Observatory. It consists of three identical 300 mm f/3 telescopes, equipped with Sloan-like g' , r' , and i' band filters, and designed to simulate a single node of the SiTian array. The terminal instrument used is an ASI6200 CMOS camera from ASI, which maintains the same basic CMOS structure as the cameras in the SiTian array. This configuration provides a field of view (FoV) of $2.^\circ 29 \times 1.^\circ 53$, with a pixel resolution of 9576×6388 , corresponding to 0.862 per pixel. For more detailed information about the telescope, refer to He et al. (2025), and for further details about the CMOS cameras, refer to Zhang et al. (2025).

During operations, the full width at half maximum (FWHM) of images typically ranges between 2' and 4'. For all fields except those designated for gravitational wave events, the Mini-SiTian telescopes operate in staring mode, consistently targeting the same patch of the sky. Differential images are generated by subtracting these observations from a high signal-to-noise ratio (SNR) reference image, which is compiled by stacking quality-filtered historical observations of the same region. This approach allows for the identification of sources that exhibit potential photometric variations. When a source is identified as real—i.e., not an artifact caused by cosmic rays or instrumental errors—based on specific criteria or algorithms, and it is not listed in any known variable star database, it is classified as a potential transient event. An alert is then triggered, prompting follow-up observations with larger telescopes to further investigate the phenomenon.

For transient search tasks within a fixed sky area, such as detecting electromagnetic counterparts of gravitational wave events where prior information about the sky region is available, the Mini-SiTian Array will adopt a scanning mode. In this mode, the array observes a segment of sky for a few exposures before moving on to the next segment. Due to the large combined field of view (FOV) of the three-color nodes in the SiTian project, it is possible to generate reference images of the entire sky relatively quickly. Consequently, the method used for transient searches in scanning mode is similar to that in staring mode. However, for the Mini-SiTian project, which represents only a single node of the SiTian array, the FoV of a single telescope is much smaller (approximately 3 deg^2 for a Mini-SiTian telescope, compared to about 25 deg^2 for a SiTian telescope).

Generating a reference image for the entire sky is not feasible during the test phase and would require roughly one year to complete. As a result, this work focuses on evaluating the performance of the staring mode rather than testing transient search tasks in the scanning mode.

This section presents a flowchart outlining the STRIP pipeline for transient detection (see Figure 1 [Figure 1: see original paper]) and discusses the rationale behind the design choices made for this pipeline.

2.1. Image Processing

Prior to daily observations, calibration frames, including flats and biases, are obtained. Preprocessing begins by stacking the bias and flat-field images from the day, resulting in the creation of a master bias and a master flat for calibration purposes. This step ensures uniform background levels and consistent pixel response across images. Without it, polynomial fitting of the background during the subsequent HOTPANTS image subtraction process can lead to uneven background levels, which may negatively affect star detection by SExtractor.

Once generated, these master frames are archived in a dedicated repository. In cases where adverse weather conditions or other factors prevent the capture of flat and bias frames on a given day, the pipeline retrieves previously archived master calibration frames to replace the missing ones. Since the characteristics of flat-field and bias frames for the CMOS cameras used in the Mini-SiTian Array do not change rapidly over time (Zhang et al. 2025), recent calibration frames are generally sufficient for rapid processing and transient source detection. Consequently, the pipeline automatically retrieves the most chronologically recent master calibration frames from the repository for calibration.

Upon completion of observations, the images are transmitted to the server hosting the STRIP pipeline, which first verifies the integrity of the received data. Following successful transmission, the STRIP pipeline automatically performs preprocessing steps, including bias-field and flat-field corrections. After preprocessing, the source extraction process is carried out using SExtractor (Bertin & Arnouts 1996), with sources detected at an SNR exceeding two sigma above the background. Images containing more than 100 stars detected by SExtractor proceed to have their World Coordinate System (WCS) information solved using “astrometry.net” (Lang et al. 2010). In cases of poor weather conditions or sparse star fields, where fewer than 100 stars are detected, the WCS solution may fail. In such instances, the affected images are excluded from further processing.

The calibrated images with updated WCS information possess a certain level of astrometric precision. However, the error associated with this astrometric solution typically remains several pixels, which does not meet the precision requirements for accurate image subtraction. Therefore, further refinement of the astrometric solution is necessary before storing the images in the good image repository. To achieve this, we first use the derived WCS information to extract

bright stars (brighter than 18.5 mag) from the Gaia catalog in the nearby sky area, saving them in LDAC format. We then employ SCAMP (Bertin 2006) to perform a more precise WCS refinement using the LDAC catalog. This refined WCS is subsequently used by SWarp (Bertin 2010) to resample the image onto the coordinate grid defined by the template image. If this image is the first to be obtained for a given target or field, its WCS serves as the template WCS for that field, and the image is stored as the initial template in the template library for alignment purposes. All subsequent images are aligned to this template WCS.

After resampling and alignment, the images are archived in a repository designated for good images. Once stored, the pipeline uses SExtractor to extract star catalogs and compares them with the Gaia DR3 catalog to calculate the magnitude zero-point and the overall average FWHM. Additionally, the limiting magnitudes are estimated. This information is used to identify images with smaller seeing and higher limiting magnitudes, which are then selected for stacking to produce a template image with an improved SNR.

The HOTPANTS image subtraction method involves convolving the template image with a series of kernels to match its point-spread function (PSF) as closely as possible to that of the science image prior to subtraction. This convolution process generally increases the FWHM of the stellar profiles. To achieve a PSF in the template image that closely matches that of the science image, the FWHM of the template should be smaller than or equal to that of the science image. Therefore, images with smaller seeing are preferred for use as the template.

Image subtraction is essentially a process in which the noise in the resulting image is the sum of the noise from both the template and the science images. To minimize noise in the subtracted image, it is crucial to obtain high SNR images by stacking multiple images with smaller seeing. Consequently, to achieve high SNR and minimized seeing, the template image should be updated regularly as new high-quality images become available in the image library. The template library will be maintained and updated through the survey proceeding. At the conclusion of each observation day, the pipeline automatically reviews the repository of high-quality images, selecting images with an FWHM of less than 4.5 pixels (approximately 3.9). When more than 20 such images are available, the STRIP pipeline will select the top 75% of these images, based on their limiting magnitude, for stacking. The resulting stacked image will then replace the old template and be stored as the updated template in the library. These selection criteria are adjusted according to the telescope's pixel size and the site's seeing conditions to ensure that images with smaller seeing are selected, while maintaining a sufficient number of images for analysis.

This process encapsulates the comprehensive image preprocessing workflow. With preprocessed observational images and regularly updated templates, the staring mode can effectively identify variable stars within the observed fields, as further detailed in Section 2.2.

2.2. Staring Mode Observations of Transient Sources

The staring mode of the Mini-SiTian project is designed to detect fast-evolving transients, such as flares that last from minutes to hours, as well as the photometric variations of supernovae that evolve over days to months. To identify these transient sources, the STRIP pipeline utilizes image subtraction software within a CPU environment. After preprocessing and image alignment using SWarp, the images are prepared for comparison with the template image using the HOTPANTS software. By inputting both the template and the resampled science image into HOTPANTS, a residual image is generated, highlighting primarily variable sources. An example of the template image, science image, and residual image is shown in Figure 2 [Figure 2: see original paper].

Given the inherent instability of image subtraction methods and the potential for cosmic rays and saturated pixels to appear in the subtraction images, the STRIP pipeline employs a real-bogus machine learning classifier to refine the source extraction process. This approach, similar to that used in ZTF (Duev et al. 2019), helps mitigate these issues. The training and operation of the classifier are described in Section 4. Since SExtractor can only detect positive-valued sources, a second source extraction process is conducted after the residual map is inverted. The real-bogus classification is then repeated, and the intersection of positive and negative detections forms the candidate variable source catalog for the observation.

This candidate catalog is archived and cross-matched with previous candidate catalogs for the same field, using a specified radius of 3 to refine the list. The coordinates from this cross-match serve as the basis for fixed-aperture photometry on both the template and subtracted images, yielding magnitude measurements that are recorded in a light curve database. Promising candidates are then prioritized for follow-up observations with larger telescopes. The flexibility of the STRIP pipeline allows for the integration of alternative filters, such as cross-matching with galaxy catalogs to identify potential supernova candidates, highlighting the dynamic and adaptable nature of the pipeline in the context of transient astronomy.

3. Operational Testing

This section presents an empirical evaluation of the STRIP pipeline’s performance in detecting two key transient phenomena—stellar flares and supernovae—during staring mode observations, utilizing both simulated data sets and real observational data collected during Mini-SiTian test runs.

3.1. Flare Triggering Simulation Test

Stellar flares are one of the most common transient sources observed in the Mini-SiTian survey and represent a primary scientific objective during its test runs. During these test runs, upon detecting a flare, the STRIP pipeline triggers

a response, initiating follow-up observations with larger telescopes. Therefore, ensuring the consistent and accurate detection of stellar flares by the STRIP pipeline is crucial. However, due to the limited number of flare events observed during testing, a statistical analysis to evaluate the pipeline's performance was not feasible.

To assess the pipeline's capabilities across a range of SNR of stars and flare intensities, we generated simulated observation images based on real Mini-SiTian data. These simulations closely replicate actual observational conditions (see Figure 3 [Figure 3: see original paper]) and feature stars exhibiting flares of varying intensities. By incorporating these simulated images into the standard observational workflow of the STRIP pipeline, we evaluated its flare detection rate and overall efficiency. As part of our workflow, a dedicated real-bogus classifier (Shi et al. 2025) is used to effectively filter out false detections. Once the classifier removes a significant number of bogus objects, the performance of the test is primarily influenced by the recall rate. Consequently, the focus of this evaluation was on measuring the recall of flare detections, with precision not considered in this particular analysis.

Our simulation involved the creation of two distinct images: one containing flare events and the other without, replicating actual observational conditions. To simulate flare events, we augmented the flux of stars in the second image. These simulated images were then processed through the STRIP pipeline. The pipeline first generates a subtracted image and subsequently searches for variable sources, identifying flare stars. By comparing the number of flare stars detected by the STRIP pipeline to the total number introduced, we assessed the pipeline's capabilities.

The process for generating simulated observational images is outlined as follows:

1. A precise flux map of stars and background within F02 was obtained by stacking nearly 100 high-quality images. The gain of the images was taken into account to calculate the precise flux at each pixel.
2. A bright star with a high SNR, extracted from the field after background subtraction, was used as a template to simulate other stars.
3. This template star was scaled by varying factors to produce host stars with SNRs of 0.5, 2.5, 10, 33, 100, 300, 1000, and 2000. Additionally, the peak brightness of these stars was increased by 30, 60, 90, 120, 180, 240, 360, and 480 ADU to simulate flare intensities. These stars were then placed in empty regions of the image to generate new flux maps.
4. To account for Poisson noise, we randomly generated noisy observational data sets based on the flux map, incorporating the flare stars. This process was repeated 100 times for each pre- and post-flare image pair. Each pair was processed through the image subtraction and source-finding routines of the STRIP pipeline. The detection rate of flare stars was then determined for fixed host star SNR-flare intensity combinations. This experiment also allowed for the evaluation of the pipeline's ability to detect supernovae, testing its effectiveness in galaxies with varying surface

brightness.

Figure 4 [Figure 4: see original paper] illustrates the detection rates for various host star SNRs and flare intensities in a single observation. The x-axis represents the SNR of the host star where the flare occurred, while the y-axis represents the increase in the peak brightness of the host star in ADU during the flare event. A fitted distribution of these detection rates provides an estimate of the STRIP pipeline's capability to detect flares, offering valuable insights for future observations. The figure reveals two regions of low detectability: one at the bottom and one on the right. The bottom region corresponds to flare intensities too faint to be detected in the subtraction image. The right region is saturated because the stars are so bright that their signal exceeds the upper limit of the CMOS detector.

3.2. Supernova/TDE Observation Trigger Test

Supernovae, as non-recurring transients with unpredictable locations prior to their explosion, remain poorly understood, particularly in terms of their early physical processes. To address this knowledge gap, the Mini-SiTian project adopts a staring mode, continuously monitoring a designated patch of sky and enabling rapid response to detect supernovae as early as possible. To evaluate whether the STRIP pipeline fulfills these objectives, this section analyzes data from 87 supernova fields observed during the test run, with particular emphasis on several fields where supernovae or TDEs occurred during the observations.

Between 2023 November 6 and 2024 June 27, 87 fields were observed between 40 and 261 times. Among these, 55 events, distributed across 39 fields, are listed on the TNS website. A closer examination revealed that five events were located near the edge of the images and subsequently fell outside the field of view in later observations after the supernova explosion occurred. Of the remaining 50 events, 12 have been confirmed as supernovae on the TNS website, while 38 remain classified as unclassified astronomical transients (ATs). The average observation intervals for these events range from 2.27 to 5.03 days, with an overall average interval of 3.40 days.

A comprehensive comparison of all 50 events within the supernova fields was conducted. Of the 12 confirmed supernovae, 2 were automatically detected by the STRIP pipeline, while 3 out of 38 unclassified astronomical transients (ATs) were identified by the pipeline. Additionally, the program observed a re-brightening event of a tidal disruption event (TDE 2020afhd). A detailed visual inspection revealed three primary reasons for the failure to detect many of these events:

1. **Small Aperture of the Mini-SiTian Telescope.** The Mini-SiTian telescope has a relatively small aperture of only 30 cm, significantly smaller than most other time-domain surveys such as ATLAS and ZTF. Combined with an average observation interval of 3.4 days, this resulted in some targets being observable at peak luminosity but becoming too faint for

detection after several days of brightness decline. Visual inspection showed that only two supernovae had faint signatures in the science images, while the others were too dim to be detected.

- 2. Limitations of the STRIP Pipeline’s Template Images.** The STRIP pipeline generates template images by stacking 20 individual images with good seeing conditions. However, during the early stages of the observation process, when fewer than 20 images are available, the pipeline must rely on single images with relatively poor SNR as templates. This limitation explains why two supernovae, initially visible in the science images, became undetectable after image subtraction. Additionally, early images may not have had optimal seeing conditions. If the science images had better seeing than the template images, the image subtraction process performed suboptimally, leading to an increased number of false detections. However, once the SiTian project reaches full operational capacity, it will conduct sky surveys every 30 minutes, enabling the collection of 20 high-quality images within a few days. This operational enhancement will address the issues encountered during the Mini-SiTian test run.
- 3. Imaging Issues Hindering Source Subtraction.** Some sources could not be effectively subtracted due to imaging issues, such as focusing problems or tracking inaccuracies, which hindered the detection of supernovae.

Figure 5 [Figure 5: see original paper] presents a comparison of the images of supernova 2024hrs and TDE 2020afhd, both before and after the image subtraction process within the STRIP pipeline. The corresponding light curve, shown as blue scatter points in the figure, was measured by the STRIP pipeline at the same field. The green dashed lines indicate the times when the supernova and TDE triggered an alarm within the pipeline. For the supernova event SN2024hrs, it is evident the STRIP pipeline triggered the alarm approximately two days earlier than the earliest report on the TNS. However, it is important to note that the program was not fully operational at the time of the supernova explosion. This early detection was realized only during post-completion testing of the program.

3.3. Program Execution Efficiency Test

The testing of supernova fields involved a comprehensive evaluation of the entire STRIP pipeline workflow, simultaneously assessing its operational efficiency. Figure 6 [Figure 6: see original paper] presents the processing times for a total of 582 images captured across 13 supernova fields during the test run. The median processing time was 119 s, with 90% of tasks completed within 152 s. Currently, the algorithm does not utilize multi-core parallel computing or GPU acceleration; however, it already meets the requirement of responding within 10 minutes. Looking ahead, with the anticipated implementation of more efficient algorithms, we expect processing speeds to improve significantly, potentially increasing several-fold.

4. Data Products During the Test Run

During the test run, in addition to observing supernovae to validate our ability to detect long-timescale transient sources, we selected three sky fields with mid-night culmination near the zenith to assess the performance of prolonged continuous staring observations. These fields, sequenced according to their observation schedules, were designated as field-01 (F01), field-02 (F02), and field-03 (F03). Observations in these fields were conducted using both the Mini-SiTian-02 telescope (SDSS-g band) and the Mini-SiTian-03 telescope (SDSS-r band). The central coordinates and field-of-view ranges for these three fields are illustrated in Figure 7 [Figure 7: see original paper]. Specifically, F01 was observed 1956 times by the Mini-SiTian-02 telescope and 1469 times by the Mini-SiTian-03 telescope. F02 was observed 3636 times by the Mini-SiTian-02 telescope and 3554 times by the Mini-SiTian-03 telescope, while F03 was observed 412 times by the Mini-SiTian-02 telescope and 491 times by the Mini-SiTian-03 telescope.

To evaluate the variable source detection capability of the STRIP pipeline, all dual-band images from F02 were thoroughly scrutinized, with every source triggering an alert logged. In the absence of an automated light curve classification system, images containing more than 100 alerts over the course of 3500 observations were flagged as potential variable source candidates, resulting in a total of 421 objects. These candidates were subsequently manually classified through visual inspection. Ultimately, the following variable sources were identified within F02: one flare event, 49 eclipsing binaries, and 78 pulsating variables. A comparison with existing variable star catalogs, such as those from ATLAS (Tonry et al. 2018) and SIMBAD Database, revealed that nine pulsating variables were not present in these catalogs and were thus classified as new variable star candidates.

Two distinct data sets were compiled to facilitate the subsequent training of machine learning algorithms that will be integrated into the STRIP pipeline:

1. **Training Dataset for Real-Bogus Classifiers.** A data set was created by cropping 64×64 pixel sub-images centered on the positions of variable sources exhibiting significant variations (real sources) and at locations incorrectly identified as variable sources (bogus sources) in subtracted images. The training data set includes approximately 4,000 positive cases and 215,000 negative cases. A sample of 100 examples from both categories is shown in Figure 8 [Figure 8: see original paper].
2. **Light Curve Dataset for Classifier Training.** Targets that triggered more than 100 alerts were classified as variable stars, with corresponding labels assigned based on visual inspection. A total of 127 light curves with obvious periodic signals were labeled as either “pulsating stars” or “eclipsing binaries,” while 293 light curves exhibiting pure noise were labeled as “non-variable sources.” These labeled light curves now form a data set for training an automatic variable source classifier. A selection of five example light curves is presented in Figure 9 [Figure 9: see original

paper].

Asteroids passing through the observed fields may trigger single responses without generating multiple alerts. By plotting response triggers, using different colors to represent different observation times within a single figure, we can visualize a series of colorful trails that facilitate asteroid detection, independent of reliance on an asteroid database. Utilizing this method, we successfully identified 14 asteroids in field F02, which represents more than half of the 26 asteroids, brighter than magnitude 20, predicted to be present in this region according to the orbital parameters provided in the Mini-SiTian asteroid analysis paper (Liu et al. 2025). Upon further inspection, all 11 asteroids brighter than magnitude 19 were detected using this approach. This outcome demonstrates the effectiveness of rapidly identifying asteroids directly from the variable source candidate catalogs generated by the STRIP pipeline. When compared with the known asteroid database, all detected asteroids were found to be listed in the Minor Planet Center (MPC) database.

5. Limitation and Improvement

Currently, the Mini-SiTian Array operates with a single node containing three telescopes for test observations, generating three $6k \times 9k$ images every five minutes. Under the current CPU configuration, the Hotpants software performs satisfactorily, completing data processing tasks within the required timeframes. However, when scaling to the full SiTian array, which will consist of over sixty telescopes, operational challenges are expected. Each telescope unit will house four $9k \times 9k$ CMOS detectors, collectively producing data at approximately two-minute intervals. At this scale, the existing Hotpants implementation may face processing bottlenecks due to limited computational resources. Potential solutions to address these issues include optimizing Hotpants' runtime efficiency or adopting GPU-accelerated computational algorithms, such as the SFFT algorithm, which can leverage rapidly advancing parallel computing capabilities.

In parallel, the STRIP pipeline successfully generates numerous transient candidate light curves in real-time during its operation. Further processing with a machine learning classifier or large language model, such as Falco (X. Zuo et al. 2025, in preparation), will utilize the multi-band information to classify transients based on their light curves. Once integrated into the STRIP pipeline, this system will enhance the pipeline's capabilities, enabling real-time transient alerts and automated source classification.

6. Conclusion and Prospects

To benchmark the scientific objectives of the SiTian project using the Mini-SiTian telescopes, we developed the STRIP pipeline. This pipeline is specifically designed to enable near real-time detection of short-timescale transient events, such as stellar flares, within ten minutes of observation, while also handling long-duration transients like supernovae and gravitational wave electromagnetic

counterparts. Our testing has confirmed the STRIP pipeline's excellent performance in the staring mode observation. Using data from supernova and TDE fields, we validated its efficiency in detecting these transients. Additionally, data from the F02 field in the test run provided insights into the reliability of the STRIP pipeline, resulting in two data sets that will be used to train advanced machine learning modules for future integration into the pipeline.

Building on the data accumulated from the initial test run, the STRIP pipeline is now prepared to tackle further challenges in the upcoming second round of Mini-SiTian observations, scheduled for 2024 September. We anticipate continuous iterations and optimizations, with the goal of significantly advancing the scientific outcomes expected from the Mini-SiTian project.

The SiTian project is a next-generation, large-scale time-domain survey designed to build an array of over 60 optical telescopes, primarily located at observatory sites in China. This array will enable single-exposure observations of the entire northern hemisphere night sky with a cadence of only 30 minutes, capturing true color (gri) time-series data down to about 21 mag. This project is proposed and led by the National Astronomical Observatories, Chinese Academy of Sciences (NAOC). As the pathfinder for the SiTian project, the Mini-SiTian project utilizes an array of three 30 cm telescopes to simulate a single node of the full SiTian array. The Mini-SiTian has begun its survey since 2022 November. The SiTian and Mini-SiTian have been supported from the Strategic Pioneer Program of the Astronomy Large-Scale Scientific Facility, Chinese Academy of Sciences and the Science and Education Integration Funding of University of Chinese Academy of Sciences.

Y.H. acknowledges the supports from the National Key Basic R&D Program of China via 2023YFA1608303, Strategic Priority Research Program of the Chinese Academy of Sciences (XDB0550103) and the National Natural Science Foundation of China (NSFC, grant Nos. 12422303 and 12261141690). K.X. acknowledges the supports from the NSFC (grant No. 12403024), the Postdoctoral Fellowship Program of CPSF under grant No. GZB20240731, the Young Data Scientist Project of the National Astronomical Data Center, and the China Post-doctoral Science Foundation (No. 2023M743447). Z.F. acknowledges the supports from the Strategic Priority Research Program of the Chinese Academy of Sciences under grant No. XDB0550000. J.F.L. acknowledges supports from the NSFC through grant Nos. 11988101 and 11933004, and supports from the New Cornerstone Science Foundation through the New Cornerstone Investigator Program and the XPLOER PRIZE.

ORCID iDs

Hongrui Gu <https://orcid.org/0009-0007-5610-6495>

Yang Huang <https://orcid.org/0000-0003-3250-2876>

Yongkang Sun <https://orcid.org/0000-0002-3935-2666>

Kai Xiao <https://orcid.org/0000-0001-8424-1079>

Zhirui Li <https://orcid.org/0000-0002-7598-9250>
Beichuan Wang <https://orcid.org/0009-0001-5324-2631>
Zhou Fan <https://orcid.org/0000-0002-6790-2397>
Chuanjie Zheng <https://orcid.org/0009-0006-7556-8401>
Henggeng Han <https://orcid.org/0000-0003-3474-5118>
Hu Zou <https://orcid.org/0000-0002-6684-3997>
Wenxiong Li <https://orcid.org/0000-0002-0096-3523>
Hong Wu <https://orcid.org/0000-0002-4333-3994>
Jifeng Liu <https://orcid.org/0000-0002-2874-2706>

References

- Alard, C. 2000, *A&AS*, 144, 363
Alard, C., & Lupton, R. H. 1998, *ApJ*, 503, 325
Allred, J. C., Kowalski, A. F., & Carlsson, M. 2015, *ApJ*, 809, 104
Becker, A., 2015 HOTPANTS: High Order Transform of PSF ANd Template Subtraction, Astrophysics Source Code Library, ascl:1504.004
Bellm, E. C., Kulkarni, S. R., Graham, M. J., et al. 2019, *PASP*, 131, 018002
Bertin, E. 2006, in *ASP Conf. Ser.* 351, *Astronomical Data Analysis Software and Systems XV*, ed. C. Gabriel et al. (San Francisco, CA: ASP), 112
Bertin, E., 2010 SWarp: Resampling and Co-adding FITS Images Together, Astrophysics Source Code Library, ascl:1010.068
Bertin, E., & Arnouts, S. 1996, *A&AS*, 117, 393
Djorgovski, S. G., Drake, A. J., Mahabal, A. A., et al. 2011, arXiv:1102.5004
Drake, A. J., Djorgovski, S. G., Mahabal, A., et al. 2009, *ApJ*, 696, 870
Duvvuri, D. A., Mahabal, A., Masci, F. J., et al. 2019, *MNRAS*, 489, 3582
Filippenko, A. V., Li, W. D., Treffers, R. R., & Modjaz, M. 2001, in *Small Telescope Astronomy on Global Scales*, *ASP Conf. Ser. Vol. 246*, ed. B. Paczynski, W.-P. Chen, & C. Lemme (San Francisco, CA: ASP), 121
Han, X., Xiao, Y., Zhang, P., et al. 2021, *PASP*, 133, 065001
He, M., Wu, H., Liang, T., et al. 2025, *RAA*, 25, 044005
Hu, L., Wang, L., Chen, X., & Yang, J. 2022, *ApJ*, 936, 157
Huang, Yang, Liu, Jifeng, Wu, Hong, et al. 2025, *RAA*, 25, 044001
Kochanek, C. S., Shappee, B. J., Stanek, K. Z., et al. 2017, *PASP*, 129, 104502
Lang, D., Hogg, D. W., Mierle, K., Blanton, M., & Roweis, S. 2010, *AJ*, 139, 1782
Law, N. M., Kulkarni, S. R., Dekany, R. G., et al. 2009, *PASP*, 121, 1395
Li, G., Hu, M., Li, W., et al. 2024, *Natur*, 627, 754
Li, W. D., Filippenko, A. V., & Treffers, R. R. 2000, in *AIP Conf. Ser.* 522, *Cosmic Explosions: Tenth AstroPhysics Conf.*, ed. S. S. Holt & W. W. Zhang (Melville, NY: AIP), 103
Liu, J., Soria, R., Wu, X.-F., Wu, H., & Shang, Z. 2021, *AnABC*, 93, 1
Liu, X. 2019, *Electronic Proceedings of “Galactic Archaeology in the Gaia Era”*, 2019 Sexten Center for Astrophysics (SCfA), 14 (Sesto Pusteria: SCfA)
Liu, Z.-X., Gao, J., Gu, H., et al. 2025, *RAA*, 25, 044010
Shappee, B., Prieto, J., Stanek, K. Z., et al. 2014, *AAS Meeting Abstracts*, 223,

225.02

Shi, J., Gu, H., Huang, Y., Zhang, Y., & Du, P. 2025, RAA, 25, 044008

Tonry, J. L. 2011, PASP, 123, 58

Tonry, J. L., Denneau, L., Heinze, A. N., et al. 2018, PASP, 130, 064505

Zhang, J.-C., Wang, X.-F., Mo, J., et al. 2020, PASP, 132, 125001

Zhang, Y., Du, L., Hu, Y., et al. 2025, RAA, 25, 044003

Note: Figure translations are in progress. See original paper for figures.

Source: ChinaXiv — Machine translation. Verify with original.