Digital Scholarship in the Humanities, 2024, **39**, 1026–1042
https://doi.org/10.1093/llc/fqae057
Full Paper

OXFORD

From past to future: digital methods towards artefact analysis

Andrew Harris [1,*], Andrea Cremaschi [2,3], Tse Siang Lim[4], Maria De Iorio [1,2], Chong Guan Kwa[5]

[1]Department of Paediatrics, Yong Loo Lin School of Medicine, National University of Singapore, Singapore 117597, Singapore
[2] A*STAR Institute for Human Development and Potential, Singapore 117609, Singapore
[3]School of Science and Technology, IE University, Chamartín, Madrid 28006, Spain
[4]Independent Scholar, Australia
[5]S. Rajaratnam School of International Studies, Nanyang Technological University, Singapore 639798, Singapore
*Corresponding author. Andrew Harris, Postdoctoral Research Fellow in Southeast Asian Archaeology and Numismatics, Yong Loo Lin School of Medicine, National University of Singapore, 21 Lower Kent Ridge Road, Singapore 119077, Singapore. E-mail: harris72@nus.edu.sg

Abstract

Over the past two decades, Digital Humanities has transformed the landscape of humanities and social sciences, enabling advanced computational analysis and interpretation of extensive datasets. Notably, recent initiatives in Southeast Asia, particularly in Singapore, focus on categorizing and archiving historical data such as artwork, literature and, most notably archaeological artefacts. This study illustrates the profound potential of Digital Humanities through the application of statistical methods on two distinct artefact datasets. Specifically, we present the results of an automated minting study of mid-first millennium CE struck and cast 'Rising Sun' coinage from mainland Southeast Asia, while subsequently utilizing unsupervised statistical methods on 2D images of 13th–14th-century earthenware ceramics excavated from the precolonial St. Andrew's Cathedral site in central Singapore. This research offers a comparative assessment showcasing the transformative impact of statistics-based approaches on the interpretation and analysis of diverse archaeological materials and within Digital Humanities overall.

Keywords: numismatics; ceramic reconstruction; unsupervised statistical methods; clustering; Southeast Asian archaeology; digital humanities.

1. Introduction

The past two decades have seen an unprecedented data-driven revolution within the study of humanities and social sciences through computer-based analysis and the application of digital technology (Schriebman, Siemens, and Unsworth 2004; Berry 2011). Known collectively as 'Digital Humanities', this field is broadly defined by Drucker as encompassing 'work at the intersection of computational methodology and humanities materials', and has been developed to better address humanistic questions, approaches, and data analysis (Drucker 2021). Originally known as 'humanities computing', Digital Humanities began as a series of early initiatives of data digitization through the creation of textual archives and databases, but has evolved to encompass methods of data interpretation such as 'computer-based statistical analysis, search and retrieval, topic modelling, and data visualisation' (Berry 2019). Given the recent advances in statistics and computer sciences, new tools are constantly being developed, each groundbreaking in the quantity of data that can be analysed and the quality of analysis that is performed (alongside efficiency and cost-effectiveness) that could otherwise not be achieved through manual, human study (Eslami et al. 2021; Karl et al. 2022; Mara 2022). Digital methods can be developed gradually and build in complexity, and can be supervised or automated, allowing scholars to focus their time and effort instead towards the creation of novel methodologies and interpretation of their data rather than exclusively data analysis. Digital Humanities also allows for substantial collections of data to be assessed through simplified workflows, for instance through Drucker's model which streamlines data into digitization (Materials), computational analysis (Processing), and publication (Presentation), striking a balance between ease of use and critical understanding (Drucker 2021). Finally, the demand for publicly available data has led to methods which emphasize open interpretation and accessibility, creating 'interpretive materials, the curation and documentation of objects, and the examination of the digital reception of heritage, particularly through social media' (Morgan 2022: 213–214).

The use of digital methods within the study of humanities is readily applied to the various sub-fields that make up archaeology, defined as 'the study of the ancient and recent human past through material remains' (SAA, 2023). Digital archaeology thus describes methods and theory that stem from the application of computational methods to archaeological studies, with the term 'computational archaeology' readily applied to computer-based systems that 'adapt existing technological innovations for specific archaeological purposes' (Aycock 2021; see Grosman 2016). Mara (2022) notes that the use of mathematical and statistical methods in archaeology occurred quite early, at the forefront of Digital Humanities (Binford 1965; Clarke 1973). Since the 1970s, digital methods have been used in a variety of archaeological investigations, for instance mapping archaeological sites through the application of GIS, the reconstruction of ancient sites and artefacts through 2D or 3D modelling, sorting and typologizing various artefacts using unsupervised or semi-supervised statistical models, and the digitization and archiving of museum and/or private collections (Papaioannou, Karabassi, and Theoharis 2002; Rasheed and Nordin 2015; Khunti 2018; Tuno et al. 2022; Natarajan et al. 2023).

The past decade has seen the rise of Digital Humanities in Southeast Asian contexts, most notably in Singapore, through the categorization and archiving of historical datasets such as artwork, literature, artefacts, and archival texts (van Lit and Morris 2024). Many of these initiatives currently comprise the simple storage of data or the creation of public or private databases, but in select cases involve the use of digital applications in historical interpretation.

Heng, for example, utilizes keyword searches within ancient Chinese texts as a means for 'data-trawling' specific topics related to Southeast Asia's premodern history, for instance trends in trade between Southeast Asia and China during specific dynastic periods through referencing specific trade goods (Heng 2019). Moreover, Klassen et al. utilize linear-regression algorithms to predict the construction dates of temples in 9th–14th-century Angkorian Cambodia from an incomplete set of foundation inscriptions (Klassen, Weed, and Evans 2018). Computational analysis, however, has never been applied to Southeast Asian artefacts recovered from archaeological contexts, which provides a unique avenue of inquiry and the application of novel statistical methods for analysis and interpretation.

To highlight the potential of Digital Humanities in Southeast Asian archaeological studies, this article focuses on two separate datasets of artefacts from first and early-second millennium CE, both thought to represent patterns of trade and exchange across Southeast Asia's 'Silk Road of the Sea' (Miksic 2013). The first is a series of silver coins minted in the mid-first millennium CE featuring a uniform but locally varied 'Rising Sun' motif. Although this motif likely originated in the Indianized polities and city-states of central and southern Myanmar, so-called 'Pyu-Mon' polities, Rising Sun coins have been excavated across mainland and peninsular Thailand and in select regions of Cambodia and Vietnam (Gutman 1978; Wicks 1992; Epinal and Gardère 2014). Using coins recently recovered from the site of Angkor Borei in Takeo Province, Cambodia and previous excavations at the important ancient port of Oc Eo, An Giang Province, Vietnam, we present the results of an automated minting study of cast and struck coins, employing recently-proposed high-dimensional clustering methods to assess whether coins found in specific and regional contexts may have been minted from the same dies or moulds, as well as to better understand the role of standardized coinage in this relatively underexplored ancient theatre of trade. The application of these methods to a dataset of coins featuring multiple production techniques is novel, as automated studies have almost exclusively been applied to die-struck specimens (Heinecke et al. 2021).

The second dataset comprises rim- and base-sherds from earthenware ceramic vessels excavated from the site of St. Andrew's Cathedral in central Singapore (Lim 2012). These sherds come from the 14th-century settlement known in historical sources as Temasek, which for a short period acted as a notable locus of international maritime trade between China and India prior to the foundation of modern Singapore in the early 19th century. We show that unsupervised classification methods using 2D images can accurately recover group of sherds belonging to the same original vessel. The article is structured as follows.

2. Minting studies and historical background

2.1 Minting studies and digital numismatics

Coin minting studies, known better as die studies, are essential within the study of ancient numismatics, and are an important tool for assessing the economic history of early states (Heinecke et al. 2021).

Ancient coins were typically minted from either hand-engraved dies, with obverse face (front) and reverse-face (back) dies struck together over annealed metal, or cast, where molten metal was poured into stone or clay moulds. In both instances, each face features unique, symbolic designs representative of familiar politico-religious symbols, or denoting a ruler themselves. These images suggest that, alongside an understood value, the initial conception and production of many currencies often reflect a specific identity ascribed to centralized governance (see Graeber 2011) Minting studies of struck, cast, or combined coins have the potential of quantifying the number of dies used in large sample sizes of coinage, including within archaeological contexts, which can also be used to compare the probable relative size of a coinage with that of another series (De Callatay 1995; Esty 2011).

From a technical point of view, die or mould identification can be seen as a classification problem with a finer subdivision of classes; for instance, every coin class consists of several different imprinting mechanisms used for striking. The appearance of an ancient coin is usually unique because no two blank coins were ever struck at exactly the same angle or with the same force and dies deteriorated over the course of the minting process, resulting in different impressions. Other factors influencing variation in coin appearance include material, craftsmanship, tools used in minting, the specific die, mint signs and shape. Furthermore, coins suffered from wear and tear, were clipped to shave off precious metal, or were marked by money changers and government authorities. Coins minted from the same die can exhibit noticeable differences due to variations in the striking process. However, coins of the same type may appear nearly identical, even if struck from different dies, because all the dies used for a particular issue were based on the same original design. Cast coins, meanwhile, can show differences based on metal composition, filling quality, cooling rate, mould wear, and post-casting treatment, for example trimming. Apart from these factors, comprehensive manual studies are an enormous time investment, and depending on the number of coins involved, are estimated to take between a year (1,000+) to a lifetime (60,000+ coins) (Aagaard and Marcher 2015; Van Alfen 2017). Furthermore, any individual coin minting study represents a regional subset of any total output, and requires contextualization within a greater body of known examples.

Digital numismatics is an approach to the study and appreciation of coins and currency exploiting computational techniques to enhance our understanding and engagement with numismatic artefacts. This field, that has received increasing attention over the last decades, combines the traditional expertise of numismatists with cutting-edge digital tools and techniques, opening new avenues for research, cataloguing, and accessibility. One of the key strengths of digital numismatics is its ability to create comprehensive and easily accessible databases of coins from various historical periods and regions. This not only aids scholars and collectors in their research endeavours, but also democratizes access to this rich cultural heritage for a broader audience. Advanced imaging technologies, such as high-resolution photography and 3D scanning, are used for detailed and accurate documentation of coins, revealing intricate details that might otherwise be overlooked (Bentkowska-Kafel and MacDonald 2018; Hess, MacDonald, and Valach 2018). This not only aids in characterization and authentication but also provides a valuable resource for the study of iconography, inscriptions, and historical context.

Referencing struck Roman coinage, Heinecke et al. (2021: 2) note that automated die studies offer greater efficiency and potentially improved accuracy in categorization compared to manual examinations, even with photographs and expert verification. In recent years, computer vision-based analysis of ancient coins has been attracting increasing attention, yet despite this research effort the results achieved remain poor and far from being useful for any practical purpose (Cooper and Arandjelović 2019). The application of computer vision methods to coin analysis has mainly focused on coin classification (van der Maaten and Postma 2006; Zaharieva, Kampel, and Zambanini 2007; Kampel and Zaharieva 2008; Anwar, Zambanini, and Kampel 2015; Sasi and Sreekumar 2015; Aslan et al. 2020). Moreover, the existing methods usually rely on the availability of a reference set, a tall order for pre-modern coinage, which comes in hundreds of thousands of distinct types. Most success has been obtained with contemporary machine-made coins (e.g., the Dagobert coin recognition system aimed at sorting high volumes of coins (Nölle et al. 2003), but algorithms for the classification of modern coins are not directly applicable to hand-minted pre-modern coins. The assumption of uniformity, and thus the straightforward feature comparison for modern coins, facilitates the classification process substantially. Most existing algorithmic approaches to coin identification are based on the extraction of local features from each coin image (Cooper and Arandjelović 2020) and on the definition of a similarity measure between pairs of coins, which is crucial for classification accuracy. These approaches often result in poor performance due to loss of spatial relationships between the different impressions on the coin. To the best of our knowledge, only a limited number of the existing proposals explicitly address the problem of die analysis, that is, the determination of the number of dies/casts/minting tools corresponding to a sample of coins and their exact partition (Natarajan et al. 2023). Moreover, current methods are mainly based on pairwise comparisons and at best use heuristics to determine the number of dies (Taylor 2020). Finally, there has thus far been no attempt to statistically estimate chronology by combining reverse and obverse information.

2.2 A brief history of first millennium CE Southeast Asian 'Rising Sun' coinage

We demonstrate the efficacy and potential of automated minting studies using an unsupervised statistical model on a sample of struck and cast coinage from first millennium CE Southeast Asia, a region featuring a relatively underexplored numismatic history compared to that of contemporary currency-based economies such as Rome, India, and Central Asia (Sutherland and Carson 1926–2019; Mitchiner 2002; Cribb and Bracey 2019). Archaeologist John Miksic (2013) among others notes that Southeast Asia formed an important midpoint for international maritime trade between India and China as early as the second millennium BC, while Roman and Chinese sources written as early as the third century BC document Southeast Asia's trade routes, ports, goods, tribute, and to a lesser extent ancient peoples and customs active in the region (Vickery 2003; Heng 2019). Miksic points out that Southeast Asians were evidently important facilitators of trade during this period, yet Mahlo argues that '[scholars] are still dealing largely with peoples, kingdoms, and ruling dynasties with unknown histories' (Mahlo 2012: 12), and thus the internal workings and dynamics of trade are often unknown or at least require further investigation through archaeological analysis.

Southeast Asia's earliest regionally produced coinages, denominations of 98 per cent pure silver, were first minted in the city-states of southern and western Myanmar as early as the fourth century CE (Gutman 1978). The introduction of coinage to Southeast Asia is not believed to have been a local innovation, instead a product of the gradual process of Indianization across Southeast Asian, which through trade and cultural exchange resulted in the integrated South Asian religion with indigenous Southeast Asian systems of local rulership and spirit-worship (Mus 1933). Apart from Arakanese coins of the Candra Dynasty (4th–10th centuries CE) and a recent find of a gold coin from southern Cambodia (Epinal and Gardère 2014: 111), Southeast Asian coinage from this period was not inscribed with the names or images of rulers. Instead, coinage featured a standardized set of symbolic Hindu-Buddhist images; while slightly augmented over time, these remained fairly uniform throughout their period of production. For instance, what are thought to be the earliest Southeast Asian coins feature the image of an auspicious conch *(sankha)* (obverse) and an aniconic Indic symbol known as a *srivatsa* (reverse) (Gutman 1978; Wicks 1992; Goyal 1995; Miksic and Goh 2017).

The most widespread type of coin produced during this period, and the focus of this study, depicts the Rising Sun (obverse) with a *srivatsa* yet again (reverse); the former features a half-risen sun surrounded by a border of 27–31 pellets, while the *srivatsa* is flanked by a swastika (Hindu-Buddhist symbol of good fortune) and *bhadrapittha* (auspicious seat) and situated below a small sun and moon of varying design (Fig. 1) (Mahlo 2012: 33–45). 'Rising Sun' coins are believed to have been predominantly manufactured and distributed between the fourth and seventh centuries CE within Southeast Asia, and feature a standard weight between 9.2 and 9.4 g with diameters ranging from 28–35 mm (Wicks 1992: 118). Local imitations vary somewhat from these measurements yet feature the same general iconography. As such, Rising Sun coins are not believed to represent any single monarch nor dynasty of this period but instead symbolize rulership in general; recounting the first millennium CE, the Burmese Glass Palace Chronicle of 1796 CE refers to the separate 'Pyu-Mon' kingdoms of Myanmar collectively as the 'Sun Dynasty' (Tin and Luce 1923: 6–7).

Rising Sun coins of a variety of subtypes have been found across mainland Southeast Asia from Arakan to southern Vietnam; however, regional groupings and relative chronologies have been established solely based on coin metrics (weights and measures) and the iconographic programmes depicted on the coins themselves (Gutman 1978; Wicks 1985). Often, the findsites of these coins are within settlements verified as important regional economic centres, which are either inland riverine settlements or coastal entrepôts (Hall 1999). Excavations at sites such as Halin, Beikthano, Maingmaw, and Sri Ksetra (Myanmar), U Thong, Nakhon Pathom, Nakhon Si Thammarat, and Lopburi (Thailand), and Oc Eo (Vietnam, see below) report numerous findings of Rising Sun coins; however, archaeological finds in Southeast Asia are often not widely reported (Mahlo 2012: 11–13; see Onwimol 2018). Nevertheless, when publicized, coin finds are either noted as single examples, often found surrounding important religious monuments, or as part of buried hoards, conglomerations of coins deposited in ceramic vessels or other containers (see Wicks 1992: 211). Hoards, which in Southeast Asia are known to number between 3 and 2000+, are incredibly important for further assessing artistic typologies beyond existing private collections (Mitchiner 2002; Mahlo 2012) as well as for determining relative minting dates, whether coinage was produced locally or imported, and even how wealth and power were perceived vis-'a-vis coinage in various regions, for example through the

Figure 1. Select results of manual die analysis.

Note: Above: Obverse and reverse Rising Sun coins from predominant cast group (ten of sixty-seven coins total shown), Angkor Borei. The coins in Figure 1 in our classification are RS0175, RS0179, RS0710, RS0717-0719, RS0722, RS0742-0743, RS0756. These correspond to SOSORO ID SSR699, SSR703, SSR15.14, SSR15.21-15.23, SSR15.26, SSR15.46-15.47, SSR15.60. Below: Matching Cast Coins (Type 8b), Angkor Borei (Upper-Left); Single Struck Specimens (Type 8a/8b), Oc Eo (Upper Right). The coins in Figure 2 in our classification are RS0210 and RS0758 (Angkor Borei, SOSORO ID N/A and SSR.17) and RS0296-0297 (Oc Eo/An Giang, HCMC ID BLTS.2206(b) and BLTS.2208(a)

Image Sources: SOSORO Museum of Economy and Money; Ho Chi Minh City Museum of History production of smaller denominations vs. the fractioning of coins (Wicks 1992: 162).

The coins analysed within this study represent the easternmost known extent of 'Rising Sun' coin finds in Mainland Southeast Asia, which in turn mark the easternmost reach of any polities which used these coins as either currency or weighted standard of silver for trade. As noted, each of the samples assessed in this study was found in the area of the Mekong Delta at two sites: the inland centre of Angkor Borei, Takeo Province, Cambodia, and within and around the ancient port-city of Oc Eo, An Giang Province, Vietnam. Both sites were inhabited contemporaneously as settlements within the so-called Kingdom of Funan (c. second–seventh centuries CE), an important early Indianized polity in Southeast Asia documented in contemporary Chinese chronicles (Vickery 2003; Stark 2004). Funan is thought to have encircled the Mekong Delta and Gulf of Thailand and provided strategic links between Indian and Chinese spheres of coastal trade (see Cœdès 1968). However, despite residing in the same cultural sphere, connected by an artificial canal (Sanderson et al. 2007), the archaeological contexts within which these coins were found differ greatly. Almost all known Rising Sun coins from Angkor Borei originate from a single hoard of struck and cast coins discovered in 2011 at Konlah Lan, near the ancient settlement's western gateway (Epinal & Gardère 2014: 108). This hoard, reported to contain over 2,000 coins, is considered one of the most significant early Southeast Asian coinage finds; however, fewer than 100 full coins, along with several fractional denominations, are currently housed in Cambodia's SOSORO Museum of Economy and Money in Phnom Penh.[1] The coins from Oc Eo and in the surrounding countryside of An Giang Province, meanwhile, were found during multiple excavations directed by French archaeologist Louis Malleret in the 1940s (Malleret 1959–1963). Unlike the Konlah Lan hoard and other finds around Angkor Borei, which primarily comprised of full-unit denominations, Oc Eo's excavations unearthed proportionately far more cut coin fragments (1/2, 1/4th, 1/8th, and 1/16th); 250 of the 262 coins samples housed at the Ho Chi Minh City (HCMC) Museum of History from Oc Eo were fragmented compared to only twenty-four recovered from the inland site (Epinal and Gardère 2014: 105). Consequently, only twelve full Rising Sun coins from Oc Eo/An Giang have been incorporated within this study, and collectively represent the total published from this archaeological region.[2]

2.3 Data

Data collection for this study was completed between November 14–17, 2022, at the SOSORO Museum of Economy and Money in Phnom Penh, Cambodia and February 21–24, 2023 at the Ho Chi Minh Museum of History, Vietnam. Coins were measured for diameter/radius and thickness, weighed, and photographed, with supplemental photographs arriving later. A test-set of ninety full-denomination Rising Sun coins (eighty-eight obverse, eighty-six reverse)[3] was included into the study, with seventy-eight coins from Angkor Borei, presumably from the Konlah Lan hoard, and twelve from Oc Eo/An Giang Province. Each coin was given a classification number beginning with RS,[4] for example, RS0100.

A comparative manual minting study of the ninety coins within this test-set was completed prior to the automatic study and used as benchmark. All coins were compared with examples from private collections and online auction catalogues; the latter have been consolidated through the Southeast Asian Numismatics Digital Archive by renowned Southeast Asian numismatic expert Robert Wicks (Wicks, *n.d.*). A larger study incorporating coins from museums in Myanmar, Thailand, and the UK is ongoing.

2.4 Pre-processing

We performed a series of pre-processing steps following Natarajan et al. (2023). Specifically, we convert the images to greyscale and resize them to 300×300 pixels. Then, we apply total-variation image restoration (Rudin, Stanley, and Emad 1992) to each image to reduce the noise introduced during image acquisition resulting from low lighting, and to even out small aberrations, while preserving the quality of edges and corners in the images. Contrast limited adaptive histogram equalization (Pizer et al. 1987) is then applied to locally enhance contrast and edge definition, while limiting noise amplification in near-constant image regions. Potential artefacts introduced by the local transformations used in the latter procedure are then reduced by a second application of total-variation restoration. See Fig. 2 for an example of raw versus enhanced coin images (obverse and reverse).

2.5 Statistical methodology

In this work, we exploit techniques from machine learning and Bayesian statistics to extract coin features and quantify differences between pairs of coins. This approach allows the specification of suitable distances to be used in distance-based clustering, with the aim of grouping the coins based on their features. In particular, we employ the D2-Net model proposed by Dusmanu et al. 2019a, b, performing dense feature extraction via a previously trained Convolutional Neural Network, namely the sixteen-layer network VGG16 proposed by Simonyan and Zisserman (2014). D2-Net yields a representation of the coin image that is

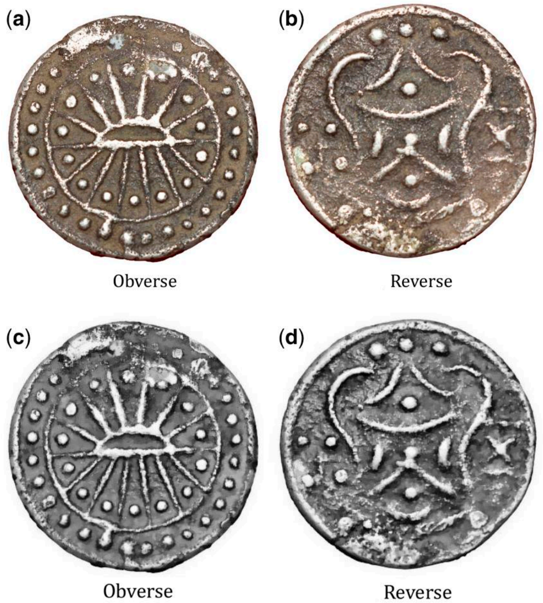

Figure 2. Example of a coin before (top row) and after (bottom row) preprocessing enhancement. Coin ID RS0175, Angkor Borei.

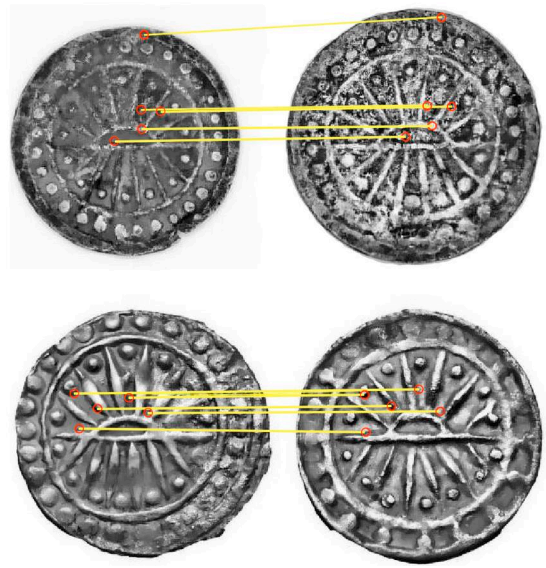

Figure 3. Example of landmark matching between two coin obverses belonging to the same (top) or different (bottom) cast/die set. All coins come from Angkor Borei (RS0175/RS0178 and RS0758/RS0761).

simultaneously a detector (i.e., a locator of keypoints in the image) and a descriptor (i.e., represented by an array of features describing the image properties at the corresponding keypoints). More details, as well as a GitHub repository with code and guidelines for implementation, can be found in Dusmanu et al., (2019a, b).

After feature extraction, we match coin image pairs via an efficient approximate nearest neighbour search (Muja and Lowe 2009). This results, for each image pair (i, j), in a subset of matched *landmarks*, selected among the D2-Net keypoints, between images i and j. We then identify the circular regions of each pair of images i and j enveloping the coins via the Matlab function *imfindcircles*. Unmatched keypoints and those located outside the coin circles are discarded. An example of the resulting landmarks identified for a pair of coins (obverse) is shown in Fig. 3. The landmark features are used to compute different dissimilarity measures between pairs of coin images. First, the matched landmarks are ranked using Gaussian Processes landmarking (Gao, Kovalsky, and Daubechies 2019a; Gao et al. 2019b), to identify the most reliable landmarks for image comparison. Thus, the features extracted by the D2-Net model corresponding to the ranked landmarks are used to compute a variety of similarity metrics between coins. These include the number of matched landmarks, weighted Euclidean distances (Natarajan et al. 2023), as well as the pairwise structural similarity index measure (SSIM), comparing the degree of similarity between two images by weighing their luminance, contrast and structure (Wang et al. 2004). Of all the possible pairs of coins to be compared, we exclude those with less than three matched landmarks. The order of the landmarks matters in the computation of some of these measures, which include rank-dependent weights. The metrics are then used to compute a final dissimilarity measure corresponding to the Euclidean norm of the vector whose components are the individual metrics. This yields a dissimilarity matrix, composed of the pairwise distances between coin pairs. Due to the exclusion criterion based on the number of matched landmarks, the dissimilarity matrix presents some missing entries, which are imputed following the ultrametric procedure (Makarenkov and Lapointe 2004).

The dissimilarity measure is employed in the distance-based clustering approach proposed by Natarajan et al. (2023). The authors, again assessing Roman coins, propose a Bayesian nonparametric model for distance-based clustering where a suitable likelihood is specified for the pairwise distances between high-dimensional objects, thus drastically reducing the curse of dimensionality. Moreover, the prior distribution on the partition has the micro-clustering property (Betancourt, Zanella, and Steorts 2022), which allows the identification of smaller, but still informative clusters. In this context, a cluster of coins represents coins minted from the same die.

The distance-based clustering model is fit to obverse and reverse images, independently and inference is performed through a Markov chain Monte Carlo (MCMC) algorithm (Natarajan et al. 2023). The obverse and reverse MCMC chains are used to obtain posterior estimates of the partition of the coins into clusters by minimizing the variation of information loss function (Meilă 2007), which combines information within each of the two partitions (entropy), as well as the information shared between the partitions (cross-entropy). The estimated clustering of the obverse and reverse coin images are formed of four and three clusters of sizes, respectively (see De Callatay 1995; van der Maaten and Postma 2006; Bentkowska-Kafel and MacDonald 2018; Aycock 2021). To assess the level of accuracy of the proposed methodology, we compute the Rand index (Rand 1971) between the estimated partitions and the results from the manual analysis. Given two partitions of the same objects, the Rand index measures the proportion of pairs that are clustered in the same way (i.e., together or separately in both partitions) over the total number of possible pairs. This index takes values in the interval (0,1), with higher values indicating a higher agreement between the two partitions under comparison. We obtain Rand indices equal to 0.77 and 0.84 for obverse and reverse, respectively.

2.6 Discussion

From our manual analysis, we concluded that of the sixty-seven coins sampled from Angkor Borei, likely originating from the Konlah Lan hoard, derive from the same minting tools (Fig. 1, Above). An additional eight, also verified by the SOSORO Museum to come from the hoard, were struck or cast from four different obverses and reverses (with a pair of coins assigned to each die/mould). The remaining three coins, each struck, featured die linkages between the obverse and reverse matches: RS0212 (O1/R1), RS0760 (O1/R2), and RS0188 (O2/R2). The twelve coins from Oc Eo, meanwhile, comprise entirely single struck specimens despite similar artistry (see Fig. 1, Below for an example). The result from the automatic data analysis is generally in line with those obtained by the manual analysis for both reverses and obverses. A summary of the comparison between the two analyses is displayed in Figs. 3 and 4. We notice that the automatic analysis can correctly recover the mix between the reverse and obverse sides as discussed above.

Although our study shows that automatic minting analysis tools still need to be perfected, from a numismatic point of view the availability of the automatic results allows for a faster sorting of coins in advance of a critical examination. The reduction in necessary comparisons results in saving numerous hours of manual labour. This is primarily because, on average, the remaining comparisons within each cluster are far less time-consuming than those required in a brute-force die or mould study. A significant portion of images is already correctly assigned, and many accurate assignments have been made. In practical terms, visual validation mainly involves making corrections for low-quality images or poorly preserved coin images. In summary, automatic minting analysis, illustrated through struck and cast coins, empowers researchers to dedicate more time to the numismatic and art-historical analysis of their material. Additionally, it provides valuable information about the minting process through die combination diagrams (Esty 1986) and offers insights into mint output via statistical analysis.

Historically, this study lays the groundwork for long-overdue examinations and scrutiny of previous typologies of Rising Sun coins, which in turn helps numismatists better examine the significance of different coin types and subtypes at various archaeological sites in mainland Southeast Asia. For instance, comparisons with coins typologies generated from the private collections of Mitchiner (12002) and to a greater extent Mahlo (2012) suggests that the variations of coins sorted from Angkor Borei and the Konlah Lan hoard collectively represent Type 8b in Mahlo's typology. The single specimens from Oc Eo, primarily sorted into Clusters 2–4, are a combination of subtypes 8a, 8b, and 8c (Mahlo 2012: 33). Mahlo's typology, however, is based on the downgrading of

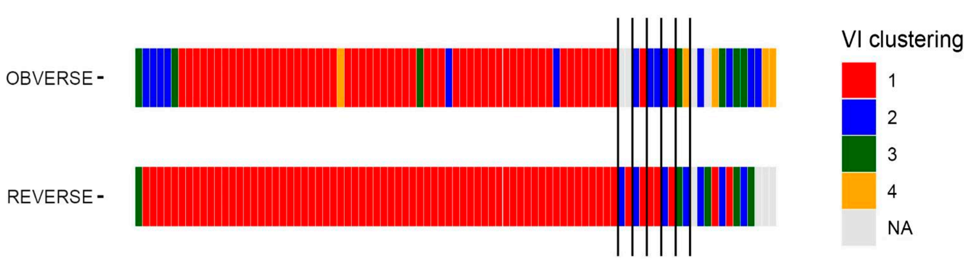

Figure 4. Clustering assignment obtained with the automatic analysis. Each element of the grid represents one of the ninety distinct coins analysed. The coins are displayed by obverse/reverse (top/bottom row), and each colour represents a different cluster. The black vertical lines separate coins belonging to different dies as by manual analysis, with the rightmost coins representing the group of single specimens. Grey grid elements indicate coins for which only one face was available in the dataset.

technical quality and 'errors' in the figuration from an 'idealised' example (8a) minted in Myanmar, rather than a detailed analysis of coins and their archaeological find-sites. Each of the six obverse-reverse die pairs (as well as the three O1/O2 links), for instance, suggest that these broad 8a, b, and c typologies indeed can be broken down into more detailed, localized subtypes connected to geography and regional activity. It is also probable that the 8a coins from Oc Eo were imported, which suggests that the Funanese port played host to trade from 'Pyu-Mon' city-states and beyond. Coins found at Angkor Borei in the Konlah Lan hoard, meanwhile, were probably locally-minted imitations used for trade (see below), rather than a centrally issued currency with politico-religious significance. Instead, we theorize these local imitations signified wealth through association with a contemporary yet overall undocumented transregional Southeast Asian maritime economy (Moore 2009; Onwimol 2018).

The results of this study also provide new insights into the methods of production of 'Pyu-Mon' currency in Funan-based settlements, an area of underexplored numismatic research, particularly in Cambodia. In reference to the sixty-seven identically matched coins, the results of both the manual and subsequent automatic minting studies have allowed for the re-evaluation of many specimens within the Konlah Lan hoard as cast rather than struck (Fig. 2), a revelation which indeed prompted us to revise our terminology from exclusively 'automated die study' to 'automated minting study'. Casting has often been seen as atypical within the manufacture of early Southeast Asian silver coinage, but is indeed known from other contemporary regional contexts; several ceramic moulds for casting 'Pyu-Mon'-type coins, both Rising Sun and earlier conch/*Srivatsa* coins, are held at Bangkok's Coin Museum and the King Narai National Museum in Lopburi, respectively (Onwimol 2018: 72). The presence of these surviving moulds sheds light on diverse trends of local imitation of coinage for use within the same economic sphere; die and cast matches from Angkor Borei, for example, suggest some replication of the volume of coins being imported into proximate Oc Eo, where no matches were found. Furthermore, the linkage between die-struck coins RS0188, RS0212, and RS0760 from Angkor Borei indicates that several series of coins, likely produced by both striking and casting, were minted in Funan beyond the numismatic influence of 'Pyu-Mon' centres. The order in which these three coins were struck may be verified in the future through studies of die wear and re-cutting, questions which are beyond the scope of this article.

Finally, pertinent to the study of Angkor Borei, Epinal and Gardère (2014: 118) suggests that the Konlah Lan hoard may have been buried in a period of political instability during the seventh century, one which, documented in Chinese sources, saw the eclipse of Funan and the rise of a polity or polities known as 'Zhenla' (see Vickery 1998). Zhenla's prominence, like the later dominance of Burmese Pagan (849–1297 CE) and Cambodian Angkor (801–1431 CE), saw to the formation of insular, redistributive agrarian states across Southeast Asia. This centralization resulted in reduced coinage output and use in the Mekong Delta region as regional trade between these polities dwindled. Thus, although no dateable material has survived to verify the time-frame of deposit of the Konlah Lan hoard, this interpretation suggests the possibility that the 8b cast subvariants, especially those sixty-seven matching coins primarily from Cluster 1, were some of the last iterations of Rising Sun coins minted in first millennium CE Cambodia.

3. Ceramics

3.1 Statistical analysis of ceramics

Ceramics (pottery) are an artefact class defined as any shaped object made of fired clay, and represent arguably the most numerous artefacts collected in archaeological investigations. The most common archaeological ceramics are typically ceramic vessels, which are typically broken into fragments known as sherds that appear either as surface-finds surrounding areas of occupation or within excavation trenches in strata corresponding with occupation. Ceramics are typically fired at low-temperature (earthenware) or high-temperature (stoneware), and are sometimes made from specially sourced clay (for example kaolin porcelain). The remains of vessels such as pots, dishes, and vases provide significant information about the daily lives of people in the past, and their study represent an important inquiry in archaeology; Karl notes that 'pottery analysis in archaeology addresses many topics ranging from the resources of the potter's clay, the forming of pottery, the vessel shapes and painting styles—including their development over time—to its use, trade, discard and reuse' (Karl et al. 2022: 195). The sherds which survive from any of these items, however, are often in a poor state of preservation, either broken or worn-off beyond recognition, and archaeologists rarely find ceramics in an optimal state for direct visual examination. Furthermore, ancient potteries are almost always excavated with missing parts, and both analysing and reconstructing these vessels manually is time-consuming and often expensive (Mara 2022).

Given the significance of ceramic study to the archaeological field, it is no surprise that many of the digital archaeological methods generated over the past two decades have been applied to reconstructing pottery objects from either sorted or arbitrary collections of excavated sherds (Kampel and Sablatnig 2003a, b; Rasheed and Nordin 2015; Eslami et al. 2020, 2021). Papaioannou Karabassi, and Theoharis (2002) were the first to propose methods for automatic three-dimensional (3D) construction based on the alignment of parts focusing on surface geometry, whereby matching was done directly through the broken edges between two arbitrary fragments utilizing optimization algorithms. Eslami et al. (2020) note that the majority of subsequent digital ceramic reconstruction methods have relied on the procurement of these types of 3D images of sherds, either through the stretching of images or 3D modelling (for example through photogrammetry or laser scanned images) to incorporate all possible geometric features. Kampel and Sablatnig (2003a, b), meanwhile, introduced an automated archival, classification, and reconstruction system using the front and back of single fragments, and automatically extracted the features of what was suggested to be the complete vessel by computing the profile in relation to the documented measurements and/or the ratios of the dimensions of the object (see Karl et al. 2022).

Despite the time-saving and cost-effectiveness of 3D image innovations compared to manual ceramic sorting and reconstruction, Eslami et al. suggest further improvement, noting that 3D reconstruction methods form a 'bottleneck in the automatic analysis of large quantities of fragments. From direct experience, the process to get a valid discrete geometric model requires, for each sherd, a mean time of about 30 minutes' (Eslami et al. 2021: 78). Instead, Eslami et al. (2021) recommend a shift towards the use of 2D images of front profiles and edge curves. This method was previously utilized by Richter et al. for reconstructing torn documents to recover texts, as well as by Brown et al. for restoring prehistorical frescoes found in excavations at the site of Thera, Greece where information such as shape, colour, texture, and roughness was recorded and utilized alongside photographs and measurements (Brown et al. 2008; Richter, Ries, and Lienhart 2011). As such, recent evaluations of the field of digital ceramic reconstruction emphasize that advances in automated analysis through statistical methods might be more effective for 2D ceramic reconstruction. In our application in Section 3.4, we followed the 2D approach.

3.2 Historical background

The ceramics analysed in this study come from excavations conducted surrounding St. Andrew's Cathedral in central Singapore, an area of ancient settlement predating the foundation of modern Singapore in the early-19th century. Over the last 40 years, large amounts of artefacts dating between the 13th and 14th centuries CE have been recovered from several archaeological sites on Singapore Island (see Miksic 2013). These sites are all located within an area bound by the city-state's Fort Canning Hill to the north, the Singapore River to the west Stamford Road to the east and the Straits of Singapore to the south. Coincidentally, Malay oral tradition and written records, most notably the Malay Annals (Sejarah Melayu (SM)), claim that a polity known as 'Singapura' was established on the island of 'Temasek' by a semi-divine prince around the same region and period. From as early as 1819, these artefacts have been associated by historians and scholars alike with a hypothetical settlement and polity derived from inferences drawn from the SM about 'Temasek-Singapura'. It is the present consensus that these artefacts belong to a complex port-city which once existed on Singapore Island in precolonial times, which is validated by both the SM and the 14th century *Description of the Barbarians of the Isles* written by Chinese voyager Wang Dayuan in 1349 (Rockhill 1914). Despite almost two centuries of both historical and archaeological research however, very few of the political, cultural and socio-economic aspects of this settlement, its relationship with neighbouring polities and its significance within Southeast Asian precolonial history can be confirmed by verifiable historical evidence. In turn, despite the notable innovations in Digital Humanities seen in Singapore in recent years (see Heng 2019), archaeology as an academic field currently remains underdeveloped despite its promising inception in the 1980s.

These ceramics were excavated from the St. Andrew's Cathedral site (STA) between 2003 and 2004, which at the time comprised the 'most extensive, longest-lasting, as well as the most systematically designed and executed archaeological project in the history of Singapore' (Miksic and Lim 2003). A team of archaeologists led by Professor John Miksic of the National University of Singapore initiated an archaeological project to recover any artefacts at that site before they were lost in the redevelopment of the site into a two level basement extension of the cathedral in December 2003 (QPP 2003). Over a period of 28 weeks between September 2003 and June 2004, the team surveyed and excavated the site within three consecutive phases: augering, test-pit excavation and salvage excavation. The entire site occupied an area of approximately 2,400 square feet (223 m^2)—northwest of the Cathedral, along Singapore's North Bridge Road—and a total of 190 excavation units measuring 2x1 m were completed by the end of the project and excavated either down to the point of artefact sterility (absence of artefacts) or the level of the water table (180cmbd[5]) (Lim 2012).

The scale of the project yielded proportionally significant results: over 330,000 pieces of artefacts from the precolonial to colonial periods, estimated to weigh around 1 ton (0.91 metric ton), were recovered from 1,009 metric tons (or 636 m^3) of excavated soil. This amount was estimated to equal 'the amount of artefacts recovered in all previous archaeological research in Singapore' (Miksic and Lim 2003) up to 2003. In other words, the STA site had an average artefact density of 519 artefacts per cubic metre; of this amount at least half can be ascribed to the Temasek-Singapura period.

It is clear from the stratigraphic soil profiles of each trench and artefact yields alongside historical sources that three phases of human occupation exist on the STA site: colonial (1819–1965), post-colonial (1965-), and precolonial (c. 13th–15th centuries), the former representing activity related to the ancient settlement of Temasek-Singapura. The artefacts found within this early layer (95–175 cmbd in each trench) primarily comprised two types of ceramic sherds: imported Chinese porcelain and stoneware dating between the 11th and 14th centuries (Lim 2012) and locally made and imported earthenware vessels; the former form the focus of this study (see below). Chinese coins minted from the Song (960–1279 CE) and Yuan dynasties (1279–1368 CE), glass beads and bangles, worked stone, shellfish and bone remains, and an impressive record of charcoal are included in this precolonial artefact assemblage (Miksic and Lim 2003).

3.3 Methodology: unsupervised clustering of image features

Ceramic sherd data are provided via images of the fragments of ceramic sets stored at the Institute of Southeast Asian Studies (ISEAS) of the National University of Singapore, where they were grouped by visual inspection by the third author. In this work, we analyse five different ceramic sets composed of a varying number of pieces, as reported in Table 1. The images are pre-processed following the same steps used in the numismatic application (see Section 2.4), with the only difference that the images are now resized to 1000×1000 pixels. The aim of this analysis is to extract features form the available images and exploit them to recover the known grouping into sets. While this is a simple task to perform with the ceramic sets at hand, it might pose as a challenging endeavour when a large number of images are to be analysed simultaneously.

After pre-processing, the enhanced greyscale images are analysed using feature detection and extraction methods, available within Matlab's *Computer Vision Toolbox*. We use the Maximally Stable Extremal Regions (MSER) algorithm (Matas et al. 2002) to detect relevant features from each image. MSER extracts as features the connected components of the sets of

Table 1. Images of ceramics sherds within each set.

Set#	Sherds
Set1	
Set2	
Set3	
Set4	
Set5 back	
Set5 front	

pixels in the image characterized by similar intensity. The algorithm only selects the stable regions among the connected components, by increasing the region size and checking that the incremental differences are smaller than a threshold provided by the user. The centres of these regions are then used for feature extraction. This process can be visualized as follows. First, an additional third dimension is added to the image representing each pixel's intensity, yielding a 3D image object. Thus, we perform a 'flooding' of the image with water by steady increments. At each step, the regions connected by water represent the connected components of the image that, if stable enough, can be extracted as features. Examples of connected regions and extracted features identified by the MSER algorithm are shown in Fig. 5 for some of the ceramic sherds available. In this work, we consider only the top 10 per cent strongest features of each image for further analysis.

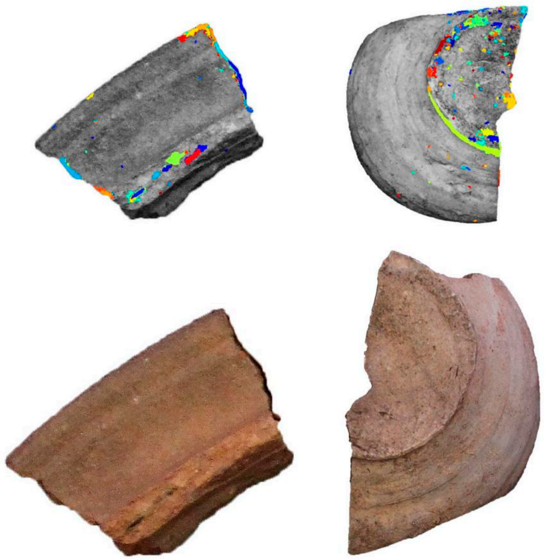

Figure 5. Top row: MSER connected components identified as circles. Bottom row: top 10 per cent strongest features extracted by MSER for further analysis.

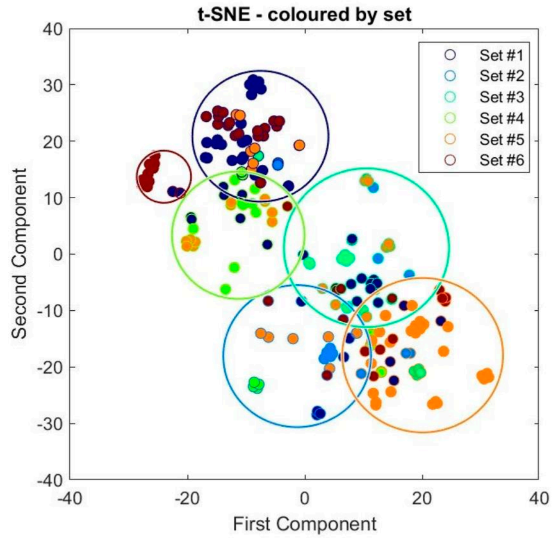

Figure 6. t-SNE projection of the features extracted from the ceramic images into a 2D space. For each point representing a feature, the inner colour refers to the available information regarding the corresponding image (i.e., which ceramic set or which ceramic sherd the feature belongs to), while the outer shape colour refers to the result of the *k*-means clustering algorithm. The circles in each image contain all the features belonging to the same *k*-means cluster.

As a result of the feature extraction procedure, we obtain a dataset of 213×64 features to use for clustering purposes. The dimension of the feature matrix results from standard feature extraction techniques. The next step in the analysis is to identify groups of images using the information contained in the feature matrix. To do so, we resort to the popular *k*-means algorithm (MacQueen 1967) and compare the result with the information arising from visual inspection. Due to the high dimension of the problem, we perform an initial dimension reduction step to improve the performance of the clustering algorithm. Specifically, we opt for the t-Stochastic Neighbour Embedding (t-SNE) approach (Van der Maaten and Hinton 2008). t-SNE is a dimension reduction technique that estimates the data's optimal projection in two or three dimensions. This technique is widely used in high-dimensional problems, such as single-cell studies, to visualize patterns within the data in a lower-dimensional setting. We apply the t-SNE approach directly to the feature matrix to obtain a 2D representation of the feature vectors. Finally, we apply the *k*-means algorithm to cluster the projected features. The number of clusters in the *k*-means algorithm is pre-specified by the user, and here we fix it to the number of observed ceramic sets, i.e. 5. We show in Fig. 6 the results of the cluster analysis and compare them with the information available regarding the ceramic images. We plot the features extracted from the ceramic images and projected into a 2D space via t-SNE. Each point represents a row of the feature matrix after t-SNE projection. Some of the images are characterized by features that are in close proximity in the t-SNE space and are therefore more easily clustered together. On the other hand, some images' features are located far apart and are therefore assigned to different clusters. Nonetheless, the method achieves an accuracy, as measured by the Rand index, of 0.7029 when compared to the set information.

3.4 Methodology: vessel reconstruction

In this section, we describe an automatic method to match sherds with the ultimate goal of reconstructing a ceramic object. We exploit computer vision methods which allow to retrace the contour of each sherd and represent it as a 2D image. Ceramic reconstruction though computer images has received increasing interest in recent years, with the development of several computational strategies in support of historical research.[6] More recently, studies by Rasheed and Nordin (2015) and Eslami et al. (2021) proposed a curve fitting approach to detect matching borders in images of broken ceramics. The 2D images are obtained through high-definition photographs taken along the vertical axis of each sherd, that is, on the side where the pottery is broken. The shape of these edges is then extracted by exploiting the Canny filter algorithm (Ding and Goshtasby 2001), a method used for border detection in binary 2D images. The outline of the shapes, typically represented as a closed ellipsoid line, is then split into two parts, each representing a different curve to be analysed. Either a polynomial curve or wavelets are fitted to these curves with the goal of both noise removal and main feature extraction. Finally, matching of the curves, and therefore of the edges of the sherds, is performed by comparing the resulting curve fittings.

In this work, we opted for a different approach, inspired by solving a jigsaw puzzle. This is motivated by the fact that available 2D images were taken perpendicularly with respect to the breaking point of the sherds, thus making it impossible to apply the methods described above. We proceeded as follows. We first applied the Canny filter for the detection of the contour of each sherd image. Then, for each pair of images within a set, we compared segments of the corresponding contours (sub-contours) by using a moving window of pixels. Each pair of sub-contours was compared following an approach analogous to that found at the GitHub repository https://github.com/MaximTerleev/Jigsaw-Puzzle-AI. In particular, we computed the seven invariants based on Hu moments (Hu 1962) of each sub-contour. These quantities describe important image features and are invariant to translation, scale, and rotation. We used them to produce different dissimilarity measures between each pair of sub-contours, as suggested by https://learnopencv.com/shape-matching-using-hu-moments-c-python/.

Differently from previous approaches, we also included the seventh invariant, which provides information about mirror images. By minimizing the sum of the squared dissimilarities for each pair of sub-contours, we were able to identify the most promising matches for where two sherds might align. More in details, for each pair of sherds, we inspected the top three sub-contour matches, and were able to correctly identify pairs of sherds to be matched together as well as the exact location on the edge. See Fig. 7 for some examples of automatic vessel reconstruction.

3.5 Discussion

Automatic methods for analysing and reconstructing ceramic sherds allow for a more comprehensive and wider-scale interpretation of the types of ceramic objects constructed and uses within early societies, especially in the absence of ancient literary sources describing these cultures in great detail. The establishment of a digital open-access archive of ceramic artefacts, too, can facilitate future both intra- and inter-site comparisons of ceramic artefacts, which would supplement existing cross-cultural and contemporary ethnographic studies. For instance, various studies chronicling ceramic production across modern mainland Southeast Asia highlight the use of either a paddle and anvil or a spatula to smoothen and shape a vessel into its final pre-firing form; a 'slow' wheel, intermittently spun by hand to shape the vessel body, may be used in the shaping process as well, but not in the same manner as the faster potter's wheel (Lefferts and Court 2003). The absence of rilling (decoration comprising fine incised close-set lines) on local coarse and medium-tempered earthenware ceramics excavated from the STA site suggests a similar practice occurred in 14th-century Temasek, which has been suggested as a product of precolonial Singaporean potters of a Malay-Nusantao or Malay-speaking people (Lim 2012). A database of reconstructed or semi-reconstructed ceramics, in this case earthenware vessels from the STA site, would help supplement or challenge these theories while also providing valuable information on the daily lives of early Singapore's residents, the island's ethnic diversity, and the longevity of various methods of ceramic production across Southeast Asia.

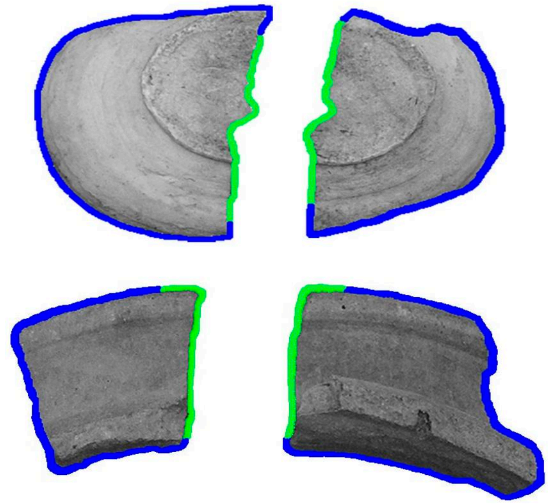

Figure 7. Matched sherd pairs obtained through automatic vessel reconstruction. Green lines represent matched edges.

Moreover, computer vision tools can be used to recover and typologize different decorative motifs, which could be, for example, used to reconstruct the art-historical evolution of specific civilizations, enabling also a faster and more comprehensive comparison across cultures.

4. Conclusions

This work highlights the increasing potential of statistical methods in Digital Humanities when applied to archaeological datasets, underscoring archaeology's position at the forefront of technological innovation within the realms of humanities and social sciences. Rather than portraying digital archaeology as deceptively straightforward, normative, de-skilling, or automating, it should be seen as an opportunity to foster meaningful engagement. These studies, it is important to note, supplement existing knowledge bases, and do not serve to substitute the hard work undertaken to contextualize the datasets of artefacts involved (Morgan 2022).

Digital Humanities represent a remarkable fusion of tradition and technology, revolutionizing, for instance, the way we approach the study and appreciation of coins and ceramics. As this field continues to evolve, it holds promise to further our understanding of history, culture, and economics through the lens of ancient artefacts. However, it is important to note that, while digital humanities holds great potential for discovery and research, there exist challenges to be overcome, such as standardization of data formats, the need for robust metadata protocols and issues related to the authenticity and integrity of digital records (Gruber and Meadows 2021). Emerging automated learning-based techniques, meanwhile, provide a valuable tool for interpreting the wealth of archaeological image-based data (photographs, 3D scans, or even hand-drawings) accumulated over more than a century of interpretation.

Digital methods of analysis allow for the enhancement of various tasks in archaeological analysis while opening the door to broader research possibilities. However, making data accessible and suitable for these methods often necessitates labour-intensive and costly processes of data structuring and annotation, which in turn are still subordinated by more traditional methods of artefact analysis, for example manual drawing and sorting. Therefore, a key requirement for the future of digital artefact analysis is the development of more efficient methods for computer-assisted preparation and annotation of large datasets of artefacts. The recent studies highlighted in this survey have set the stage for pursuing such objectives in the future (Karl et al. 2022).

Acknowledgements

We would like to thank Dr. Blaise Kilian and associated staff of the SOSORO Museum of Economy and Money in Phnom Penh, Cambodia, as well as Dr. Hoang Anh Tuan and associated staff from the Ho Chi Minh City Museum of History, for permissions to examine, measure, and photograph all 'Rising Sun' coins used in this study. Special thanks to Dr. Tran Ky Phuong for helping attain permissions in Vietnam, and for Dr. Robert Wicks for his initial assessment of our methodology and field collection methods.

Author contributions

Andrew Harris (Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Writing—original draft), Andrea Cremaschi (Conceptualization, Investigation, Methodology, Resources, Software, Writing—original draft), Tse Siang Lim (Data curation, Formal analysis, Writing—original draft), Maria De Iorio (Conceptualization, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Supervision, Writing—original draft, Writing—review & editing), Chong Guan Kwa (Conceptualization, Writing—review & editing)

Funding

This article was written underGrant No. MOE-T2EP40121-0021 from the National University of Singapore.

Data availability

The data underlying this article will be shared on reasonable request to the corresponding author.

Notes

1. Although unpublished originally, the provenance of the 67 matched coins in this study (Angkor Borei, and thus likely the Konlah Lah hoard) has subsequently been confirmed by one of the coauthors of the original museum summary (Gardère, *pers. comm.*).
2. Four others have been recovered from 'Funanese' contexts: two from an undocumented site outside of Ho Chi Minh City (Saigon) in 1886, and two in Cambodia without further specificity. The current whereabouts of both pairs of coins are unknown (Cappon 1886).
3. Due to issues with accessing the Oc Eo Rising Sun coins, only one face from six of the coins (4 obverse, 2 reverse) are featured in the study.
4. 'Rising Sun', abbreviated.
5. Centimetres below datum, an arbitrary xyz coordinate set by the excavators of any archaeological trench.
6. See Eslami et al. 2020 for an extensive review on the topic.

References

Aagaard, S., and Marcher, M. (2015) 'The Microscope Drawing Tube Method (MTDM)—An Easy and Effective Way to Make Large-Scale Die Studies', *The Numismatic Chronicle*, **175**: 249–62.

Anwar, H., Zambanini, S., and Kampel, M. (2015) 'Coarse-Grained Ancient Coin Classification Using Image-Based Reverse Side Motif Recognition', *Machine Vision and Applications*, **26**: 295–304.

Aslan, S., Vascon, S., and Pelillo, M. (2020) 'Two Sides of the Same Coin: Improved Ancient Coin Classification Using Graph Transduction Games', *Pattern Recognition Letters*, **131**: 158–65.

Aycock, J. (2021) 'The Coming Tsunami of Digital Artefacts', *Antiquity*, **95**: 1584–9.

Bentkowska-Kafel, A., and MacDonald, L. (2018) *Digital Techniques for Documenting and Preserving Cultural Heritage*, p. 370. Leeds, UK: Arc Humanities Press.

Berry, D. M. (2011) 'The Computational Turn: Thinking about the Digital Humanities', *Culture Machine*, **12**: 1–22.
Berry, D. M. (2019) 'What are Digital Humanities?', The British Academy Blog, https://www.thebritishacademy.ac.uk/blog/what-are-digitalhumanities/, accessed 9 Nov. 2023.
Betancourt, B., Zanella, G., and Steorts, R. C. (2022) 'Random Partition Models for Microclustering Tasks', *Journal of the American Statistical Association*, **117**: 1215–27.
Binford, L. R. (1965) 'Archaeological Systems and the Study of Culture Process', *American Antiquity*, **31**: 203–10.
Brown, B. J. *et al.* (2008) 'A System for High-Volume Acquisition and Matching of Fresco Fragments: Reassembling Theran Wall Paintings', *ACM Transactions on Graphs*, **27**: 1–9.
Cappon, P. (1886) 'Trouvaille de Monnaies en Cochinchine', *Revue Numismatique*, **3**: 295–7.
Clarke, D. (1973) 'Archaeology: The Loss of Innocence', *Antiquity*, **47**: 6–18.
Cœdès, G. (1968) *The Indianized States of Southeast Asia*. M. Smithies (trans). Honolulu, HI: University of Hawaii Press.
Cooper, J., and Arandjelović, O. (2019) Understanding ancient coin images. In INNS Big Data and Deep Learning conference. Springer, Cham, pp. 330–40.
Cooper, J., and Arandjelović, O. (2020) 'Learning to Describe: A New Approach to Computer Vision Based Ancient Coin Analysis', *Sci*, **2**: 27.
Cribb, J., and Bracey, R. (2019) *Kushan Coins: A Catalogue Based on the Kushan, Kushano-Sasanian and Kidarite Hun Coins in The British Museum, 1st-5th Centuries AD*. London: British Museum Press.
De Callatay, E. (1995) 'Calculating Ancient Coin Production: Seeking a Balance', *The Numismatic Chronicle*, **155**: 289–311.
Ding, L., and Goshtasby, A. (2001) 'On the Canny Edge Detector', *Pattern Recognition*, **34**:721–5.
Dusmanu, M. *et al.* (2019a) D2-Net: A trainable CNN for joint detection and description of local features. arXiv: 1905.03561 [cs.CV]. https://arxiv.org/abs/1905.03561
Dusmanu, M. *et al.* (2019b) 'D2-Net: A Trainable CNN for Joint Detection and Description of Local Features', in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 8092–8101. https://doi.org/10.1109/CVPR.2019.00829.
Drucker, J. (2021) *The Digital Humanities Coursebook: An Introduction to Digital Methods for Research and Scholarship*. New York, NY: Routledge.
Epinal, G. and Gardère, J. D. (2014) *Cambodia from Funan to Chenla: A Thousand Years of Monetary History*, pp. 93–125. Phnom Penh: SOSORO Museum of Money and Economy.
Eslami, D. *et al.* (2020) 'Review of Computer-Based Methods for Archaeological Ceramic Sherds Reconstruction', *Virtual Archaeology Review*, **11**: 34–49.
Eslami, D. *et al.* (2021) 'A Semi Automatic Reconstruction of Archaeological Pottery Fragments from 2D Images Using Wavelet Transformation', *Heritage*, **4**: 76–90.
Esty, W. (2011) *The Geometric Model for Estimating the Number of Dies*. Bari, Italiy: Edipuglia, 43–58.
Esty, W. W. (1986) 'Estimation of the Size of a Coinage: A Survey and Comparison of Methods', *The Numismatic Chronicle (1966-)*, **146**: 185–215
Gao, T., Kovalsky, S. Z., and Daubechies, I. (2019a) 'Gaussian Process Landmarking on Manifolds', *SIAM Journal on Mathematics of Data Science*, **1**: 208–236.
Gao, T. *et al.* (2019b) 'Gaussian Process Landmarking for Three-Dimensional Geometric Morphometrics', *SIAM Journal on Mathematics of Data Science*, **1**: 237–67.
Goyal, S. R. (1995) *The Dynastic Coins of Ancient India*. Jodhpur, India: Kusumanjali Prakashan.
Graeber, D. (2011) *Debt: The First 5000 Years*, pp. 320–1. Brooklyn, NY: Melville House.
Grosman, L. (2016) 'Reaching the Point of No Return: The Computational Revolution in Archaeology', *Annual Review in Anthropology*, **45**: 129–45.
Gruber, E. and Meadows, A. (2021) *Numismatics and Linked Open Data*. New York: ISAW Papers, 16. https://doi.org/10.26262/ISAW.PAPERS.V16.7017
Gutman, P. (1978) 'The Ancient Coinage of Southeast Asia', *Journal of the Siam Society*, **66**: 8–21
Hall, K. R. (1999) 'Coinage, Trade, and Economy in Early South Asia and its Southeast Asian Neighbours', *The Indian Economic and Social History Review*, **36**: 431–59.
Heinecke, A. *et al.* (2021) Unsupervised statistical learning for die analysis in ancient numismatics. arXiv, 2112.00290v.1 [es.CV] 1 Dec 2021. https://arxiv.org/abs/2112.00290v1
Heng, D. (2019) 'Premodern Island-Southeast Asian History in the Digital Age: Opportunities and Challenges through Chinese Textual Database Research', *Bijragen Tot De Taal-, Land- En Volkenkunde*, **175**: 29–57.
Hess, M., MacDonald, L. W., and Valach, J. (2018) 'Application of Multi-Modal 2D and 3D Imaging and Analytical Techniques to Document and Examine Coins: An Example of Two Roman Silver Denarii', *Heritage Science*, **6**: 1–22.
Hu, M. K. (1962) 'Visual Pattern Recognition by Moment Invariants', *IRE Transactions on Information Theory*, **8**:179–87.
Kampel, M., and Sablatnig, R. (2003a) 'An Automated Pottery Archival and Reconstruction System', *Journal of Visualization and Computer Animation*, **14**: 111–20.
Kampel, M., and Sablatnig, R. (2003b) 'Rule Based System for Archaeological Pottery Classification', *Pattern Recognition Letters*, **28**: 740–7.
Kampel, M., and Zaharieva, M. (2008) 'Recognizing Ancient Coins based on Local Features', in *International Symposium on Visual Computing*, pp. 11–22. Berlin and Heidelberg: Springer.
Karl, S. *et al.* (2022) 'Advances in Digital Pottery Analysis', *IT—Information Technology*, **64**: 195–216. https://doi.org/10.1515/itit-2022-0006
Klassen, S., Weed, J., and Evans, D. (2018) 'Semi-Supervised Machine Learning Approaches for Predicting the Chronology of Archaeological Sites: A Case Study of Temples from Medieval Angkor, Cambodia', *PLOS One*, **13**: e0205649. https://doi.org/10.1371/journal.pone.0205649
Khunti, R. (2018) 'The Problem with Printing Palmyra: Exploring the Ethics of using 3D Printing Technology to Reconstruct Heritage', *Studies of Digital Heritage*, **2**: 1–12.
Lefferts, L., and Cort, L. A. (2003) 'A Preliminary Cultural Geography of Contemporary Village-based Earthenware Production in Mainland Southeast Asia', in Miksic J. (ed.)

Earthenware in Southeast Asia, pp. 300–10. Singapore: Singapore University Press.
Lim, T. S. (2012) 14th Century Singapore: The Temasek Paradigm. Thesis submitted for the title of Master of Arts, Department of Southeast Asian Studies, National University of Singapore.
Mahlo, D. (2012) *The Early Coins of Myanmar (Burma): Messengers from the Past – First Millennium AD*. Bangkok: White Lotus Press.
Makarenkov, V., and Lapointe, F. J. (2004) 'A Weighted Least-Squares Approach for Inferring Phylogenies from Incomplete Distance Matrices', *Bioinformatics*, **20**: 2113–21.
Malleret, L. (1959–1963) *L'archéologie du Delta du M'ekong*, 3 vols, pp. 134–7. Paris: EFEO.
Mara, H. (2022) 'Digital Archaeology', *IT—Information Technology*, **64**: 193–4. https://doi.org/10.1515/itit-2022-0061
Matas, J. *et al.* (2002) 'Robust Wide Baseline Stereo from Maximally Stable Extremal Regions', in Proc. BMVC, London, UK, 2002. https://doi.org/10.5244/C.16.24
MacQueen, J. (1967) 'Some Methods for Classification and Analysis of Multivariate Observations', in *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, Vol. 1, No. 14, pp. 281–97. Berkeley, CA: University of California Press.
Meilă, M. (2007) 'Comparing Clusterings—An Information Based Distance', *Journal of Multivariate Analysis*, **98**: 873–95. MR2325412. https://doi.org/10.1016/j.jmva.2006.11.013.
Miksic, J. (2013) *Singapore and the Silk Road of the Sea*. Singapore: NUS Press.
Miksic, J., and Goh, G. Y. (2017) *Ancient Southeast Asia*. New York: Routledge.
Miksic, J., and Lim, T. S. (2003) 'Archaeological Research on The Padang and in the St. Andrew's Cathedral Churchyard: St. Andrew's Cathedral Archaeological Research Project Progress Report Summary September 2003–June 2004', *ARI Working Paper Series*, 3. Singapore: Archaeological Research Institute, National University of Singapore.
Mitchiner, M. (2002) *Ancient Trade and Early Coinage*. London: Cromwell Press.
Moore, E. (2009) 'Archaeology of the Shan Plateau: The Bronze to Buddhist Transition', *Contemporary Buddhism*, **10**: 91–110.
Morgan, C. (2022) 'Current Digital Archaeology', *Annual Review of Anthropology*, **51**: 213–31. DOI: https://doiorg.libproxy1.nus.edu.sg/10.1146/annurev-anthro-041320-114101
Muja, M., and Lowe, D. G. (2009) 'Fast Approximate Nearest Neighbors with Automatic Algorithm Configuration', in *International Conference on Computer Vision Theory and Applications*. VISAPP. Lisbon, Portugal: INSTICC Press, pp. 331–40.
Mus, P. (1933) 'Cultes Indiens et Indigènes au Champa: L'Inde Préaryenne et l'Asie des Moussons, La Religion Védique et le Brahmanisme, La Synthèse Hindouiste, Formes Actuelles des Cultes Chams, Les Kut, Le Culte Cham des Linga, Survivances et Profondeur de l'Influence Indienne au Champa', *BEFEO*, **33**: 367–411.
Natarajan, A. *et al.* (2023) 'Cohesion and Repulsion in Bayesian Distance Clustering', *Journal of the American Statistical Association*, 1–11.
Nölle, M. *et al.* (2003) 'Dagobert—A New Coin Recognition and Sorting System', in *Proceedings of the 7th International Conference on Digital Image Computing-Techniques and Applications* (DICTA'03), Sydney, Australia.
Onwimol, W. (2018) Coinage in Thailand During 4th-11th Century AD. A Thesis Submitted in Partial Fulfilment of the Requirements for Master of Arts (Archaeology), Department of Archaeology Graduate School, Silpakorn University Academic Year 2018 Copyright of Graduate School, Silpakorn University.
QPP (2003) *Pamphlet: St. Andrew's Cathedral's Quiet Places Project. A Project to Construct an Extension to the Existing Church Building*. Singapore: The QPP Building Committee, St Andrew's Cathedral.
Papaioannou, G., Karabassi, E., and Theoharis, T. (2002) 'Reconstruction of Three-Dimensional Objects through Matching of Their Parts', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **24**: 114–24.
Pizer, S. M. *et al.* (1987) 'Adaptive Histogram Equalization and Its Variations', *Computer Vision, Graphics, and Image Processing*, **39**: 355–68.
Rand, W. M. (1971) 'Objective Criteria for the Evaluation of Clustering Methods', *Journal of the American Statistical Association*, **66**: 846–50.
Rasheed, N. A., and Nordin, M. J. (2015) 'A Survey of Computer Methods in Reconstruction of 3D Archaeological Pottery Analysis', *International Journal of Advanced Research*, **3**: 712–24.
Richter, F., Ries, C. X., and Lienhart, R. (2011) 'A Graph Algorithmic Framework for the Assembly of Shredded Documents', in *Proceedings of the 2011 IEEE International Conference on Multimedia and Expo*, pp. 1–6. Barcelona, Spain: IEEE.
Rockhill, W. W. (1914). 'Notes on the Relations and Trade of China with the Eastern Archipelago and the Coast of the Indian Ocean During the Fourteenth Century', *T'oung Pao*, **15**: 419–47.
Rudin, L. I., Stanley, O., and Emad, F. (1992) 'Nonlinear Total Variation Based Noise Removal Algorithms', *Physica D: Nonlinear Phenomena*, **60**: 259–68.
Sanderson, D. C. W. *et al.* (2007) 'Luminescence Dating of Canal Sediments from Angkor Borei, Mekong Delta, Southern Cambodia', *Quarternary Geochronology*, **2**: 322329.
Society of American Archaeologists (SAA) (2023). 'What is Archaeology', https://www.saa.org/about-archaeology/what-is-archaeology, accessed 12 May 2023.
Sasi, A., and Sreekumar, K. (2015) 'Image Based Coin Recognition System: A Survey', *International Journal of Computer Applications*, **131**(1): 1–6. https://doi.org/10.5120/ijca2015906258
Schriebman, S., Siemens, R., and Unsworth, J. (2004) *A Companion to Digital Humanities*. Malden, MA: Blackwell Publishing.
Simonyan, K., and Zisserman, A. (2014) Very deep convolutional networks for large-scale image recognition. arXiv preprint, arXiv:1409.1556.
Stark, M. T. (2004) 'Pre-Angkorian and Angkorian Cambodia', in P. Bellwood and I. Glover (eds) *Southeast Asia: From Prehistory to History*, pp. 89–119. Oxford: Routledge & Curzon.
Sutherland, C. H. V., and Carson, R. A. G. (eds) (1926–2019) *Roman Imperial Coinage*, **10** vols. London: Spink and Sons.
Taylor, Z. M. (2020) 'The Computer-Aided Die Study (CADS): A Tool for Conducting Numismatic Die Studies with

Computer Vision and Hierarchical Clustering', *Computer Science Honors Theses*. **54**. https://digitalcommons.trinity.edu/compsci honors/54

Tin, P. M., and Luce, G. H. (1923) 'The Glass Palace Chronicle of the Kings of Burma', *Journal of the Burma Research Society*, **13**: 1–60.

Tuno, N. *et al.* (2022) 'Evaluation of Handheld Scanner for Digitization of Cartographic Heritage', *Journal of Cultural Heritage*, **54**: 31–43.

Van Alfen, P. (2017) 'The Computer Aided Die Studies Program', http://numismatics.org/pocketchange/cads/

Van der Maaten, L., and Hinton, G. (2008) 'Visualizing Data using t-SNE', *Journal of Machine Learning Research*, **9**: 2579–2605.

van der Maaten, L. and Postma, E. (2006) 'Towards Automatic Coin Classification', pp. 19–26. Vienna, Vienna, Austria: Proc. of the EVA.

van Lit, L. W. C., and Morris, J. H. (2024) *Digital Humanities and Religions in Asia: An Introduction*. Berlin, Boston: De Gruyter. https://doi.org/10.1515/9783110747607

Vickery, M. (1998) *Society, Economics, and Politics in pre-Angkor Cambodia: the 7th–8th Centuries*. Tokyo: Centre for East Asian Cultural Studies for UNESCO, The Toyo Bunko.

Vickery, M. (2003) 'Funan Reviewed: Deconstructing the Ancients', *BEFEO*, **90–91**, 101–143. https://doi.org/10.3406/befeo.2003.2238

Wang, Z. *et al.* (2004) 'Image Quality Assessment: From Error Visibility to Structural Similarity', *IEEE Transactions on Image Processing*, **13**: 600–12.

Wicks, R. (1985) 'The Ancient Coinage of Mainland Southeast Asia', *Journal of Southeast Asian Studies*, **16**: 195–225.

Wicks, R. (1992) *Money, Markets and Trade in Early Southeast Asia: The Development of Indigenous Monetary Systems to AD 1400*. Ithaca, NY: Cornell University Press.

Wicks, R. [Online] 'Southeast Asian Numismatics Digital Archive (SEANDA)', https://seanda.omeka.net/

Zaharieva, M., Kampel, M., and Zambanini, S. (2007, August). 'Image Based Recognition of Ancient Coins', in *International Conference on Computer Analysis of Images and Patterns*, pp. 547–54. Berlin and Heidelberg: Springer,

Digital Scholarship in the Humanities, 2024, 39, 1026–1042
https://doi.org/10.1093/llc/fqae057
Full Paper