Colin Wilder, Sam T. McDorman, Jun Zhou, Adam King, Yuhang Lu, Karen Y. Smith, Song Wang, and W. Matthew J. Simmons[1]

Snowvision: The Promise of Algorithmic Methods in Southeastern Archaeological Research

Abstract: This article presents the contexts, methods, contributions, and preliminary findings of Snowvision, a digital archaeology project developed by faculty and students at the University of South Carolina and the South Carolina Department of Natural Resources. Snowvision uses computer vision to reconstruct southeastern Native American paddle designs from the Swift Creek period, ca. 100-850 CE. In this essay, we first present the context of the Swift Creek culture of the southeastern United States, along with broader related issues in prehistoric archaeology. Then, the relevant methods from archaeology and computer vision are introduced and discussed. We also introduce World Engraved, our public-facing digital archive of sherd designs and distributions, and explain its role in our overall project. We then explore, in some level of technical detail, the ways in which our work refines existing pattern-matching algorithms used in the field of computer vision. Finally, we discuss our accomplishments and findings to date and the possibilities for future research that Snowvision provides.

Keywords: Pottery, computer vision, Native Americans, archaeology, Southeastern United States, pattern-matching, algorithmic methods

The American archaeological record is filled with fragmentary objects of bone, pottery, shell, stone, wood, and cloth variously embellished with realistic and abstract designs. These designs include figural imagery, like that seen on ancient Mayan pottery, and the abstract carvings found on marine shell gorgets amongst the Mississippian peoples of the eastern North American Woodlands. Humanities and social science scholars have put the variety of these designs to many uses. They can be used to assign dates to objects and places, trace ancient trade routes, understand ancient creative processes, and explore how images were used to create personal and collective identities. Without question, most of these topics are best addressed using complete designs rather than design fragments. As a result, the millions of broken cultural heritage objects stored in museums remain largely unstudied from a design perspective, and large numbers of decorated objects found in the archaeological record contribute little to our understanding of style, production, use, and meaning.

Many such partial designs are today found only as intricate impressions stamped onto the surfaces of pottery sherds recovered from archaeological sites in what are now the US states of South Carolina, Tennessee, Georgia, and along

---

[1] The Snowvision project has been a collaborative effort from its inception. For transparency regarding contributions to this article and to the project described herein, we recognize the following roles, adapted from the CRediT taxonomy (Brand, et al.): Project Conceptualization, K.Y.S., S.W., and C.W.; Article Conceptualization: C.W.; Software, J.Z. and Y.L.; Investigation, J.Z., Y.L., and S.T.M.; Writing – Original Draft, S.T.M., J.Z., K.Y.S., A.K., C.W.; Writing – Review & Editing, W.M.J.S., S.T.M., C.W., and K.Y.S.; Funding Acquisition, K.Y.S., S.W., C.W., J.Z., and A.K.; Data Curation, S.T.M. and Y.L.; Supervision, S.W., K.Y.S., and J.Z.

the Gulf Coast of southeastern North America. This ornate decorative style, known amongst American archaeologists as "Swift Creek Complicated Stamp," was produced throughout southeastern North America between ca. 100-850 CE. These designs, several of which are shown in their reconstructed form in Figure 1, represent one of the most significant indigenous decorative traditions of the Americas. They were produced by artisans using wooden paddles with raised carvings to "stamp" wet clay. Many southeastern archaeologists have long recognized that complete paddle designs could be reconstructed from fragmentary impressions on potsherds. By the 1950s, these archaeologists also began to realize the remarkable fact that certain unique paddle designs were widely distributed across the region, sometimes on archaeological sites hundreds of kilometers apart (see Figure 2). Thus, by studying the spatial distribution of paddle designs, once reconstructed and mapped, archaeologists could gain unprecedented insight into the pre-colonial social networks of ancient North America.

Our project, Snowvision, is named for a pioneer in the study of Swift Creek pottery, the artist and researcher Frankie Snow. Launched in 2016, Snowvision's founding purpose was to investigate the possibilities of using computer-vision technologies to analyze fragmentary sherds in order to reconstruct, and thereby study, southeastern Native American carved wooden paddle designs and the extent of their geographical spread. Understanding how widespread these designs were will give unprecedented insight into the social and economic networks of the pre-colonial southeast.

Archaeological Context and Survey of Scholarship

Pottery is ubiquitous on many archaeological sites. This ubiquity speaks to the tremendous utility and malleability of pottery as a container for processing, storing, and consuming commodities—mainly food stuffs—across human societies. Pottery also can be shaped and modified to convey subtle aspects of social life, both individual and group, such as social status and affiliation. Studying pottery, especially how it was made and decorated, can help archaeologists understand the movements, connections, foodways, technologies, and cultures of people who existed hundreds or thousands of years ago.

Identifying *stylistically defined* types of pottery allows archaeologists to apply relative dates to sites based on the technology and decoration utilized for and on the vessels. The ceramic traditions of southeastern North America are well established in archaeological literature and are defined both by technology and decorative style. The earliest tradition, Stallings, began around 2,450 BCE and is represented by large, hand-build, fiber tempered vessels found along the Savannah River from the Coastal Zone into the Piedmont (Sassaman 400). By 450 BCE, taller, coil-built, non-fiber temper pottery was made and used virtually everywhere in the Southeast (416). The walls of coil-built vessels were thinned using an anvil stone and wooden paddle to seal the coils together, and the paddles were often carved with parallel or perpendicular lines or wrapped with cordage that left impressions on the exterior of the vessel.

Around 100 CE, the paddle carvings changed to more intricate, curvilinear designs that incorporated many kinds of shapes and symmetries—this was the beginning of Swift Creek Complicated Stamp. This type of pottery was produced throughout the southeast until 850 CE (Williams and Elliott 1). Swift Creek pottery is abundant across the present-day US states of Georgia and the Gulf Coast of Florida, and likely originated in this region (Chase 51). Swift Creek designs found on types with different temper or production techniques, including Santa Rosa-Swift Creek in the western Florida panhandle, middle Woodland Pickwick Complicated Stamped wares in Tennessee, and Mann Complicated Stamp in Indiana, demonstrate the edges of Swift Creek influence (Chase 50, 55-56; Elliott 21; Smith, B.A. 112). While complicated stamped ceramic vessels continued to be produced into the historic period by some indigenous peoples, including potters of the Cherokee and Muscogee Nations, the complexity and variety of designs is greatest during the Swift Creek period.

Modern Swift Creek research would not be possible without the paddle design reconstruction work done by Bettye J. Broyles and Frankie Snow, our project's namesake. Since the wooden paddles do not survive in the archaeological record, both artists pieced together designs from fragments of vessels, called sherds, to reconstruct the original paddle carving. The work of both artists show that one design can be found at multiple *sites*, indicating connections between settlements.[2]

The work of Broyles and Snow illustrated connections between many sites across multiple states. Broyles's work focused primarily on sites around eastern Tennessee and the Chattahoochee River bounding Georgia and Alabama, and in 1968 she published more than 80 full design reconstructions with dozens of partial design fragments that primarily focused on connections between Fairchilds Landing (9SE14), Kolomoki (9ER1), and Quartermaster (9CE42)[3]. Broyles maintained

---

[2] Archaeologists use this concept of the *site* to organize, analyze, and discuss places of past human occupation and use. Archaeological sites can represent the locations of former camps, hamlets, villages, towns, cities, farms, plantations, or any number of other human social units. Sites tend to have spatial boundaries determined on the basis of the distribution of artifacts and features, both above and below ground. Thus, archaeological sites and the data gleaned from them are inherently spatial, as human behavior is inherently spatial. Because of this, significant effort is put into recording the locational information of sites and artifacts.

To register sites' locational information, archaeologists in the United States decades ago developed a standardized, state-based system, called State Archaeological Site Files. Here, sites are given a unique identifier based on a numbering system produced by the Smithsonian Institute. This identifier is further tied to both the site's geographical location as well as a limited set of additional site attributes, the details of which vary by state. However, although site locations are an important data point for the archaeologist, the exact location of a site is often legally restricted, ethically restricted, or both, so as to protect the site from damage or plunder and to respect the culturally-sensitive nature of many sites. If the site is on federal land, the exact site location is protected information under federal law. Archaeological site locations should not be disseminated through any means other than through the state archaeological site files or through permission of the landowner.

[3] The alphanumberic codes in parentheses indicate the aforementioned state archaeological site file identifier for the named site.

hundreds more unfinished designs in her personal collection; these were scanned by the University of Georgia's Laboratory of Archaeology and generously shared with our team.

Snow's work has concentrated on central Georgia, specifically Hartford (9PU1) and the Ocmulgee-Big Bend region, and he has published a selection of designs for these areas (Snow, "An Archaeological Survey…"; "Swift Creek Design Investigations"). However, Snow has reconstructed more than 400 designs from sites across Georgia, many of which remain unpublished and therefore inaccessible to anyone outside of Swift Creek research. Snow has graciously allowed all his designs to be used in developing the Snowvision algorithm.

Building on this corpus of design work, researchers are exploring many facets of the Swift Creek period through the study of this ornate pottery. Broadly, changes in vessel production and style across time and space can be used to understand changes or regional differences in foodways, technology, and society (Sassaman). The uniqueness of Swift Creek designs allows for a more nuanced understanding of the movements and connections of the people who created the vessels. This is especially true inasmuch as studies of stylistic variability have highlighted the fact that Swift Creek designs show a level of uniqueness and creativity that was unmatched in subsequent eras (Smith, K.Y. and Knight, "Style In Swift Creek Paddle Art"; "Swift Creek Paddle Designs").

Investigation of design symmetry at the site level is well-established as a means of exploring the relationship between pottery designs and social trends. For example, Pluckhahn has analyzed differences in design symmetry between artifacts recovered from Swift Creek-era villages and the sacred earthen mounds spaces often found near them, suggesting different sociocultural valences for different designs (Pluckhahn). Wallis, meanwhile, has studied how analysis of the composition of materials—whether the source of the ceramics are local or non-local—can be used to demonstrate the presence of networks of exchange across geographical areas, possibly as gifts tied to sociocultural rituals (Wallis). Many researchers have acknowledged the need for more site-specific research on design distribution and variability to answer questions about settlements interactions, to produce more precise chronologies, and to study if changes in worldviews and/or cosmologies are reflected in ceramic technology and how those changes may have influenced groups and trade networks (Anderson 299).

Overview of Snowvision

Snowvision arose from a conversation between Swift Creek researchers Karen Y. Smith and Scot Keith; Colin Wilder, Associate Director of the University of South Carolina's Center for Digital Humanities, and Song Wang, Professor of Computer Science at UofSC, joined the collaboration soon after. The now eight-member team includes faculty, staff, and students from the fields of archaeology, digital humanities, research computing, information science, and computer vision. After four years of intensive sherd data collection and algorithm development, we have reached our first milestone: sherd-to-design matching. This work has been supported by the

United States government, via the National Center for Preservation Technology and Training, the National Science Foundation, and the National Endowment for the Humanities.[4] The goal of the current NEH grant is to link the Snowvision algorithm to World Engraved (http://www.worldengraved.org), a publicly-accessible digital repository that allows researchers to contribute sherd and design content, enabling the expansion of the database. World Engraved also allows for researchers to submit 3D scans to the backend for algorithm matching and training, helping to improve the precision and accuracy of the Snowvision algorithm.

This accomplishment has raised additional challenges. First, archaeological datasets that are publicly accessible must take legal and ethical standards related to site location into consideration. However, in order to facilitate our research goal of illustrating the distribution patterns of various paddle designs—and thus to describe the social and economic networks of the Swift Creek era—the Snowvision database needs to accommodate, store, and deliver spatial data linked to the sherds on which Swift Creek designs are found. Yet, because of legal and ethical concerns, we must to be cautious with how we display location information on World Engraved. Our solution is to only deliver site geographic locations at a gross-level via the map function on the World Engraved website. World Engraved also provides the registered site number, when that is available. The site number can be used by researchers to access specific site locations by requesting such information from the respective State Archaeological Site Files. Even if the exact site location is not part of the delivered dataset, it is still analytically profitable to know that several designs are found on a single site or that one design is found on several different sites.

Our work has used two other existing datasets to guide our metadata-collection methodology: the Digital Index of North American Archaeology (DINAA) and the Archaeometry Laboratory at the University of Missouri Research Reactor (MURR). DINAA aggregates state site file information into a publicly accessible and digestible form. Like World Engraved, DINAA protects the exact location, but makes available other site attributes. Using similar site attributes coded in similar ways ensures the discoverability of our data delivered on World Engraved by the users of DINAA, and vice versa. MURR is a research laboratory that specializes in the compositional analysis of pottery sherds and other materials from archaeological sites around the world. We based our sherd submission template in part on that of MURR, in order to facilitate interoperability among the attributes collected.[5] Connecting compositional data with design and sherd data generated by Snowvision is an important facet of the research. Our project thus improves the ability of scholars to discover paddle design matches within

---

[4] The work described herein was supported by a one-year National Center for Preservation Technology and Teaching grant awarded in 2016 and a two-year National Science Foundation grant (#1658987) awarded in 2017. In 2019, the team was awarded a two-year National Endowment for the Humanities Digital Humanities Advancement Level III grant (HAA-266472-19) to deliver Snowvision to researchers and the public through World Engraved.

[5] It is possible that some sherds in our Snowvision database may also be in MURR. We cannot be certain. Exclusivity is not an issue as far as we are aware.

and across sites. Ultimately, this helps archaeologists reconstruct how foragers organized and moved across the landscape 1,500 years ago.

Our algorithmic approach also enables us to study the position or rotation of the paddle relative to the vessel. From this, we can begin to infer the handedness of potters (Sassaman and Rudolphi; Vidal). Our approach also helps us classify paddle designs. Classification helps the archaeologist sort designs by time period. Classification also has the potential to identify individual paddle makers or communities of paddle makers that share the same stylistic conventions and, by inference, are members of the same learning or practice groups. Although classification and handedness studies could be done manually, our computational approach expedites such inquiries across a large dataset and, we contend, should remove bias in the classification process. Identifying design communities through classification and handedness studies are novel contributions of our work.

The late Bettye J. Broyles may well have been the first archaeologist to expand on the idea that certain paddle designs were widely distributed across the lower Southeast, sometimes separated by hundreds of kilometers (Broyles; Holmes; Snow "Swift Creek Designs and Distributions," "Kolomaki and Milamo," "Swift Creek Design Innovations"; Snow and Stephenson; Wallis and O'Dell). From this realization was borne interesting, novel research possibilities that Swift Creek scholars have sought to address ever since. For his part, Snow ("An Archaeological Survey") noted a high number of Swift Creek design matches among sites within the big bend area of the Ocmulgee River (known as Ocmulgee-Big Bend) and between the Ocmulgee-Big Bend and the headwaters of the Satilla River to the south. These matches occurred within an approximately 48 km radius, providing, he argued, clues to settlement patterns within the region. He also noted a more limited number of design matches between those sites and sites much further to the northwest, southwest, and southeast, well outside of the Ocmulgee-Big Bend region. These longer-distance, inter-regional, matches suggest another kind of social process at work, perhaps long-distance expeditions to larger communities of aggregation for information and resource collection.

Documenting design connections across household deposits within a village site also holds research promise and is strong motivation for studying sherds and designs. Saunders examined Swift Creek designs in individual household middens at the Kings Bay site, Georgia ("Attribute Variability"). Personal communication from 2015 shows that Snow and Stephenson studied designs among individual middens at the Hartford site, Georgia. Both Saunders and Snow and Stephenson identified paddle design connections that linked each household trash deposit to one or more deposits across their respective study sites (Saunders; Snow and Stephenson). Smith and Knight have taken these patterns to indicate that paddle production and use operated independently of pottery discard, but further data is needed to test this idea (Smith, K.Y. and Knight, "Swift Creek Paddle Designs" 127).

Using instrumental neutron activation analysis—a method of analyzing the composition of ceramics—to source the geographical provenance of clays, Wallis documented instances in which Swift Creek vessels—whole pots, not sherds—

apparently moved great distances ("The Swift Creek Gift"). Among the vessels he analyzed, Wallis found that Swift Creek vessels from mortuary mounds on the lower St Johns River, Florida, were made with non-local clays, whereas Swift Creek vessels from residential sites in the same area were made with local clays (114). Many vessels of nonlocal origin may have come from sites much further up the Atlantic Coast. Wallis's argument about the long-distance movement of pots found in mortuary contexts was bolstered by independent evidence related to paddle design matches among sites.

The Snowvision project, both in our algorithm and the use of World Engraved as data-sharing repository for archaeologists, thus complements existing research and provides new possibilities for expanding these research methods. As noted, the core of our project is a robust matching algorithm based on cutting-edge computer vision technology.

Computer Vision Methods

Computer vision research offers an array of recently developing methods which complement the efforts of archaeologists by harnessing technology to achieve archaeological goals and deal with challenges posed by the necessarily fragile nature of archaeological materials.

Computer-aided identification of the designs from fragmented cultural objects has attracted great interest among archaeologists and computer scientists in recent years (Halíř; Kampbel), and the Snowvision project develops a new framework to identify the underlying carved wooden paddles impressed on pottery from the Carolinas to the Gulf Coast. To illustrate the value of computer vision technology to Swift Creek research, let us consider a case study of the elaborately carved wooden paddles of the southeastern Woodlands and the ornate curvilinear paddle impressions on countless pottery sherds of the Swift Creek style, as shown in Figure 3.

Identifying the full curvilinear paddle design from fragmentary sherds is challenging. First, each sherd only contains a small portion of the underlying full paddle design. Second, the available sherds rarely come from the same vessel, and it is difficult to assemble them into large pieces for more complete curve patterns. Third, one carved paddle will be applied multiple times on the vessel surface with spatial overlap. As a result, curve patterns detected on sherds may be incomplete or very noisy[6] due to both the gap when applying a planar carved paddle onto a curved pottery surface and the erosion of sherd surfaces over centuries. Furthermore, a sherd may contain a composite pattern—a small fragment of multiple, partially overlapping copies of the same design, as shown in Figure 3(b). Such a composite pattern is not simply a portion of the full design.

To address the above challenges, we have used the findings of recent computer vision research to develop a new framework for identifying carved paddle designs from pottery sherds. An overview of the process is given here,

---

[6] "Noise" is here (and subsequently) meant in the sense in which it is commonly used in computer science—irrelevant or meaningless data that obscures computational analysis.

with technical details spelled out in subsequent paragraphs. Broadly speaking, we extract the curve pattern from a sherd and then match it to each known design in a database and return the best matched design. As shown in Figure 4, this is done over three iterative steps: 1) extract a curve pattern from a sherd, 2) identify the underlying design for a sherd with a non-composite (single) pattern, or 3) identify underlying design for a sherd with a composite (multiple) pattern. Building on the research of Long, Shelhamer, and Darrell, we extract curve patterns using a fully convolutional neural network (Long, et al.). These FCNN-based curve pattern segmentation methods form a digitized sherd's depth map and then match the sherd with a non-composite pattern by combining a template-matching algorithm with a dual-source convolutional neural network, building on the work of Krizhevsky, Sutskever, and Hinton (Krizhevsky, et al.). The CNN re-ranking algorithm then finds the sherd's underlying design, or matches it with a composite pattern, using a new Chamfer matching algorithm, per the methods demonstrated by Zhou, Yu, Smith, Wilder, Yu, and Wang (Zhou, et. al.).

To state the process differently: given a pottery sherd, the first step of our framework is to extract curve pattern from this sherd. Generally speaking, extracting a curve pattern from the surface of a sherd is a low-level image segmentation problem of the sort typical in computer vision research. However, erosion and sediment usually make the visibility of the curve pattern on the sherd very weak and blurred, which substantially increases the difficulty in accurately segmenting them. Early in the development of our algorithm, we used the excavated pottery sherds associated with the Woodland period Swift Creek type for experiments and found that it is very difficult to extract these curve patterns from the camera-taken images of these sherds. Given that these curved patterns are stamped on the surfaces of pottery vessels by carved paddles, these patterns usually show greater depth than the adjacent non-curved surface. Therefore, 3D scanners are usually applied to achieve the 3D depth image of the sherd surface, as illustrated in Figure 5, and the curve patterns are then segmented directly from the depth image.

However, due to erosion from being buried under the earth for centuries, together with possible shallow stamping or deliberate smoothing when making the vessel, the curve patterns can still be difficult to segment even from the 3D-scanned, high-resolution depth images. Snowvision team members developed a CNN-based algorithm to more accurately and reliably segment the stamped curve patterns from the depth images of the sherds, by learning and incorporating the implied curve geometry, such as curve smoothness and parallelism, in the underlying designs (Lu, et al.). In other words, we deal with these challenges by training an FCN to detect the skeletons of the curve patterns in the depth images. Then, we train a dense prediction convolutional network to identify and prune false positive skeleton pixels. Finally, we recover the curve width by a scale-adaptive thresholding algorithm to get the final segmentation of curve patterns. Figure 6 shows the sample results after each step of this algorithm. We also extract the boundary of a sherd, indicated by red contours in Figure 6. The sherd boundary provides a mask to exclude all the information outside the sherd boundary from matching in the later steps. This CNN-based algorithm can segment the curve pattern from a sherd much more accurately than other low-level and high-level image segmentation algorithms.

However, most of the sherds contain non-composite patterns, i.e. only one copy of partial designs is present on these sherds. The second step of our framework is thus to identify underlying designs for sherds with non-composite patterns. The segmented curves from the above step can be far from perfect because of the strong noise and shallow stampings on the unearthed sherds. In particular, the curve pattern segmented from a sherd may show deformation from its underlying design due to dehydration and shrinkage during the firing process when finishing the vessel. In this step, we elaborate on a two-stage matching algorithm that is robust to noise, errors and deformation present in the segmented curve patterns. We formulate this problem by matching the curve pattern segmented from a sherd against each location, with each possible orientation, of each known design, and then select the design with the lowest matching cost as the matched design. This exhaustive matching procedure identifies not only the matched design, but also the matched location and orientation on the matched design. Based on this problem formulation, the key issue is then the definition of an appropriate cost in matching the curve patterns segmented from a sherd to a location of a full design, with a specified orientation.

This problem is nontrivial in archaeological applications for two reasons. First, the exhaustive matching against each possible location and orientation of each design leads to a very large search space. To prevent the algorithm from slowing down too much, we require the matching cost to be very efficient to compute for each possible solution in the search space. Second, compared with the underlying design, the curve patterns segmented from the sherd usually contain strong noise and deformations from the firing process, degradation from many years spent under the earth, and the imperfectness of the curve-segmentation algorithms.

To address this problem, we developed a new two-stage matching algorithm, with a different matching cost in each stage, as shown in Figure 7. In Stage 1, we propose to use a computationally-efficient classical template matching method over the whole search space to identify a small set of candidate matchings on all the known designs. The extracted sherd mask is applied as a region of interest in matching to exclude information outside this region. This simple matching cost can help efficiently reduce the search space of solutions. In Stage 2, we further derive a new matching cost by training a dual-source Convolutional Neural Network. We then apply this more computationally-intensive matching to re-rank the candidate matchings identified in Stage 1. This CNN architecture is shown in Figure 8(a).

The CNN architecture contains two identical sub-networks, which take candidate matchings and sherd curve patterns as the inputs, respectively. Each sub-network consists of a sequence of convolution, max pooling layers and a global average pooling layer (GAP) for feature learning. We implement this dual source CNN by truncating AlexNet, a CNN developed by Krizhevsky, as shown in Figure. 8(b), to "conv4" layer and replacing all layers after "conv4" layer with a GAP layer. A new matching cost is then derived by comparing the similarities between the features learned after the GAP layers. Through this supervised learning, various kinds of noise and deformations in the segmented curve patterns can be implicitly identified and suppressed in computing the CNN-based matching cost.

In the making of the pottery, the carved paddle is stamped on the pottery multiple times to ensure full coverage of the surface. As a result, a large number of the pottery sherds contain composite patterns, i.e. each pattern is a part of its underlying design, and these patterns can overlap with each other. Classical matching methods, such as Chamfer matching, require one pattern to be a portion of the other. This is not true in this case, as the curve pattern on the sherd is a composite one. As explored in previously-presented work, Snowvision team members dealt with this issue by developing a new algorithm that can automatically identify multiple components of the composite pattern extracted from the sherd (Zhou, et al.).

Taking sherd curve pattern images and design images, we first use a standard edge-thinning algorithm to reduce the curve width to one pixel as illustrated in Figure 9 (b). Although the width of curves presents an important clue in matching a sherd and a design, we try not to use the curve width information because of the difficulty of accurately measuring the curve width from a deteriorated sherd surface. Second, we extended the classical Chamfer matching method to match the one-pixel-wide curve patterns from a sherd against each location, with each possible orientation, of each known design. Different from the classical Chamfer matching algorithm, we do not pick the design with the lowest matching cost. Instead, for each design, we select several matchings as candidates, so long as each candidate fulfills a threshold percentage of total pixel matches. Shown in Figure 10, these candidates are then combined and reconstructed. The combination with the most matching pixels (which we interpret as a marker completeness) and least overlapping pixels is taken as the best matching, and its normalized completeness is taken as its matching score. The design with the highest matching score is selected as the sherd's underlying design.

To evaluate our framework, we collected a set of 1,000 Swift Creek pottery sherds that were excavated from various archaeological sites located in southeastern North America. Of these sherds, 900 contain non-composite patterns that represent 98 unique paddle designs, and the remaining 100 contain composite patterns representing 20 unique paddle designs. Each sherd in the set only displays one design, while that same design may be applied to the surfaces of multiple sherds. We use the CMC ranking metric developed by DeCann and Ross to evaluate the matching performance, as shown in Figure 11 (DeCann and Ross). In CMC curves, the higher the better. The experiments show our framework performs much better than several state-of-art algorithms.

Findings and Accomplishments

Through both our building upon and expanding multiple state-of-the-art algorithmic pattern-matching techniques *and* the creation of World Engraved, the Snowvision project has already made several significant contributions to southeastern archaeology. Nevertheless, much remains to be learned about the cultures who produced this pottery, and we believe that Snowvision will help shape the development of future scholarship in this area. Archaeological fieldwork is ongoing, which means that new data is being collected. Our algorithm greatly reduces the time spent identifying matches, and design-to-sherd scaling issues encountered during algorithm development have shown that a fraction of reconstructed paddle designs are actually design variants reproduced on two paddles with slight differences. Interdisciplinary collaboration has made the matching algorithm possible and is being carried into other aspects of the project.

We have made fruitful use of methods and ideas from information science in building both our database and World Engraved. Metadata schemas are being developed to assist researchers in submitting rich and accurate datasets, and a user needs study is being conducted to gather feedback and opinions on the algorithm, database, and website. While these tools are well-established within information science practice and literature, they are underutilized within archaeological digitization projects. Information professionals have created, refined, and freely published dozens of metadata schemas, such as Dublin Core or VRA Core, to promote resource discovery, access, and sharing (Park & Tosaka). Instructional templates that function as schemas are created for submission of archaeological data to specialized laboratories, such as those created by the Archaeometry Laboratory at the University of Missouri Research Reactor. Our schemas go beyond these instructional templates to incorporate additional fields from published schemas, such as rights statements and date of creation, that are relevant for digital publication. User assessments are commonplace within information science to ensure digital collections are relevant to users and to incorporate their needs into new or existing information systems (Mills; Green and Courtney; Wu, et al.). Preliminary data shows that the primary use of Snowvision will be to further research on the connections and movements of populations and that it will provide an avenue for student training and collaborative research. Our experiences have also indicated that researchers need access to robust data, that we need to expand to other types of complicated ceramics outside of the Swift Creek type, and that mechanisms need to be created in order to ensure that data is properly attributed to the source.

This information-focused work is being done to ensure that the Snowvision database adequately contributes to comparative research, broad syntheses, and publication of standardized archaeological data. Large-scale synthesis is needed to bring our understanding of the past from the site level to the regional level, and American archaeologists must make a return on public investment by sharing and preserving the knowledge they gain for future generations (Altschul et al., 2018). Snowvision has accomplished goals relating to synthesis, public access, and preservation. Snowvision's complementary public digital repository, World Engraved, allows for user-driven submission of data, enabling laboratories across southeastern North America to share standardized datasets for synthesis and publication in a digital archive. This will allow users to access the data held at many scattered institutions. There is no cost for any of the online Snowvision or World Engraved services, greatly expanding the accessibility of this archaeological data from a handful of researchers to any person with an internet connection. All of the original designs drawn by Broyles were lost after her death, and the digital copies held unpublished by the University of Georgia were the only records that remained. Snowvision has preserved her designs for future use and provided public access through World Engraved.

A final word should perhaps be said about public dissemination. In a way that may be unexpected to casual readers, careful consideration must be given to how sensitive data is released. Digitization of legally restricted site information has been accomplished using the example set by DINAA, but this is only one kind of sensitive data. While Swift Creek pottery is commonly found as fragmented utilitarian cooking vessels from middens and village sites, the designs were also imprinted on vessels that have been found in mound burial contexts. The excavation, study, and display of burial objects was frequently undertaken without descendant community consent for hundreds of years prior to the US Congress's passage of the Native American Graves and Repatriation Act in 1990. This has placed archaeological artifacts into a discussion of who owns the past based on cultural and intellectual property claims, created a formalized system of repatriation for human remains and funerary objects, and opened new avenues for collaborative research (Breske). While discussion of who owns the past continues, the benefits and challenges of digitization remain largely absent from the discourse. Snowvision is working with established researchers connected to descendent communities to accommodate the inclusion of sherds and designs from burial context. We believe through consideration, respect, and conversation, Snowvision can provide access to this data in ways that benefit researchers and the public and contribute to a more equitable digital archaeological future. The result, we hope, will be a fuller—and more respectful—picture of the complex social, cultural, and economic practices and networks of ancient southeastern North America than has been previously possible.

Works Cited

Altschul, Jeffrey H., et al. "Fostering Collaborative Synthetic Research in Archaeology." *Advances in Archaeological Practice*, vol. 6, no. 1, 2018, pp. 19-29. doi: 10.1017/aap.2017.31.

Anderson, David G. "Swift Creek in a Regional Perspective." *World Engraved: Archaeology of the Swift Creek Culture*, edited by Mark Williams & Daniel T. Elliott, University of Alabama Press, 1998, pp. 274-300.

Barrow, Harry, et al. "Parametric Correspondence and Chamfermatching: Two New Techniques for Image matching." International Joint Conference on Artificial Intelligence, 1977. Conference Presentation.

Brand, Amy, et al. "Beyond Authorship: Attribution, Contribution, Collaboration, and Credit." *Learned Publishing*, vol. 28, 2015, pp. 151–155. doi:10.1087/20150211.

Breske, Ashleigh. "Politics of Repatriation: Formalizing Indigenous Repatriation Policy." *International Journal of Cultural Property*, vol. 25, 2018, pp. 347-373. doi:10.1017/S0940739118000206.

Broyles, Bettye J. "Reconstructed Designs from Swift Creek Complicated Stamped Sherds." *Southeastern Archaeological Conference Bulletin*, vol. 8, 1968, pp. 49-74.

Chase, David W. "Swift Creek: Lineage and Diffusion." *World Engraved: Archaeology of the Swift Creek Culture*, edited by Mark Williams and Daniel T. Elliott, University of Alabama Press, 1998, pp. 48-60.

DeCann, Brian and Arun Ross. "Relating ROC and CMC Curves via the Biometric Menagerie." IEEE Sixth International Conference on Biometrics: Theory, Applications and Systems, 2013. Conference Presentation.

Elliott, Daniel T. "The Northern and Eastern Expression of Swift Creek Culture: Settlement in the Tennessee and Savannah River Valleys." *World Engraved: Archaeology of the Swift Creek Culture*, edited by Mark Williams & Daniel T. Elliott, University of Alabama Press, 1998, pp. 19-35.

Green, Harriett E., and Angela Courtney. "Beyond the Scanned Image: A Needs Assessment of Scholarly Users of Digital Collections." *College & Research Libraries*, vol. 76, no. 5, 2015, pp. 690-707. doi:10.5860/crl76.5.690.

Halíř, Radim. "An Automatic Estimation of the Axis of Rotation of Fragments of Archaeological Pottery: A Multi-Step Model-Based Approach." International Conference in Central Europe on Computer Graphics, Visualization and Interactive Digital Media, 1999. Conference Presentation.

Holmes, W.H. "Caribbean Influence in the Prehistoric Art of Southern States." *American Anthropologist,* vol. 7, 1894, pp. 71–79.

Kampbel, Martin and Robert Sablatnig. "Rule Based System for Archaeological Pattern Recognition." *Pattern Recognition Letters*, vol. 28, no. 6, 2007, pp. 740-747.

Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. "Imagenet Classification with Deep Convolutional Neural Networks." Advances in Neural Information Processing Systems, 2012. Conference Presentation.

Long, Jonathan, Evan Shelhamer, and Trevor Darrell. "Fully Convolutional Networks for Semantic Segmentation." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015. Conference Presentation.

Mills, Alexandra. "User Impact on Selection, Digitization, and the Development of Digital Special Collections." *New Review of Academic Librarianship*, vol. 21, 2015, pp. 160-169. doi:10.1080/13614533.2015.1042117.

Park, Jung-ran., and Yuji Tosaka. "Metadata Creation Practices in Digital Repositories and Collections: Schemata, Selection Criteria, and Interoperability." *Information Technology & Libraries*, vol. 29, no. 3, 2010, pp. 104-116. doi: 10.6017.ital.v29i3.3136.

Pluckhahn, Thomas J. "Reflections on Paddle Stamped Pottery: Symmetry Analysis of Swift Creek Paddle Designs." *Southeastern Archaeology*, vol. 26, no. 1, 2017, pp. 1-11.

Sassaman, Kenneth E. "Woodland Ceramic Beginnings." *The Woodland Southeast*, edited by David G. Anderson and Robert C. Mainfort, Jr., University of Alabama Press, 2002, pp. 398-420.

Saunders, Rebecca. *Attribute Variability in Late Swift Creek Phase Ceramics from Kings Bay, Georgia*. University of Florida, Master's Thesis. 1986.

Smith, Betty A. "Neutron Activation Analysis of Ceramics from Mandeville and Swift Creek." *World Engraved: Archaeology of the Swift Creek Culture*, edited by Mark Williams and Daniel T. Elliott, University of Alabama Press, 1998, pp. 112-129.

Smith, Karen Y., and Vernon James Knigth, Jr. "Style in Swift Creek Paddle Art." *Southeastern Archaeology*, vol. 31, no. 2, 2012, pp. 143-156. doi:10.1179/sea.2012.31.2.002.

Smith, Karen Y., and Vernon James Knigth, Jr. "Swift Creek Paddle Designs and the Imperative to be Unique." *Southeastern Archaeology*, vol. 36, no. 2, 2017, pp. 122-130. doi:10.1080/0734578X.1246342.

Snow, Frankie. "An Archaeological Survey of the Ocmulgee Big Bend Region." *Occasional Papers from South Georgia*, Number 3. Douglas, GA: South Georgia College, 1977.

___ . "Swift Creek Design Investigations: The Hartford Case." *World Engraved: Archaeology of the Swift Creek Culture*, edited by M. Williams and D.T. Elliott, University of Alabama Press, 1998, pp.61-98.

___ . "Swift Creek Designs and Distributions: A South Georgia Study." *Early Georgia* vol. 3, 1975, pp. 38–59.

___ . "Kolomoki and Milamo–Some Synchronic Evidence." *The Profile* vol. 36, 1982, pp. 7–12.

Snow, Frankie, and Keith Stephenson. "Swift Creek Designs: A Tool for Monitoring Interaction." *World Engraved: Archaeology of the Swift Creek Culture*, edited by Mark Williams and Daniel T. Elliott, University of Alabama Press, 1998, pp. 99-111.

Stephenson, Keith, Frankie Snow, and Karen Y. Smith. "Design Connections in Early and Late Swift Creek." Southeastern Archaeological Conference 69th Annual Meeting, Baton Rouge, LA., 2012. Conference Presentation. doi:10.13140/RG.2.2.12839.78242.

Wallis, Neill J. *The Swift Creek Gift: Vessel Exchange on the Atlantic Coast*. The University of Alabama Press, 2011.

Wallis, Neill J., and Amanda O'Dell. "Swift Creek Paddle Designs from the Florida Gulf Coast: Patterns and Prospects." *The Florida Anthropologist*, vol. 64, 2011, pp. 187–205.

Williams, Mark & Daniel T. Elliott. "Swift Creek Research: History and Observations." *World Engraved: Archaeology of the Swift Creek Culture*, edited by Mark Williams and Daniel T. Elliott, University of Alabama Press, 1998, pp. 1-11.

Wu, Annie, et al. "Hitting the Road Towards Greater Digital Destination: Evaluation and testing DAMS at University of Houston Libraries." *Information Technology & Libraries*, vol. 35, no. 2, 2016, pp. 5-18. doi:10.617/ital.v35i29152.

Yuhang Lu, et al. "Curve-Structure Segmentation from Depth Maps: A CNN-

Based Approach and its Application to Exploring Cultural Heritage Objects." AAAI Conference on Artificial Intelligence, 2018. Conference Presentation.

Zhou, Jun, et al. "Identifying Designs from Incomplete, Fragmented Cultural Heritage Objects by Curve-Pattern Matching." *Journal of Electronic Imaging*, vol. 26, no. 1, 2017, pp. 11-22. doi:10.1117/1.JEI.26.1.011022.

Appendix: Figures

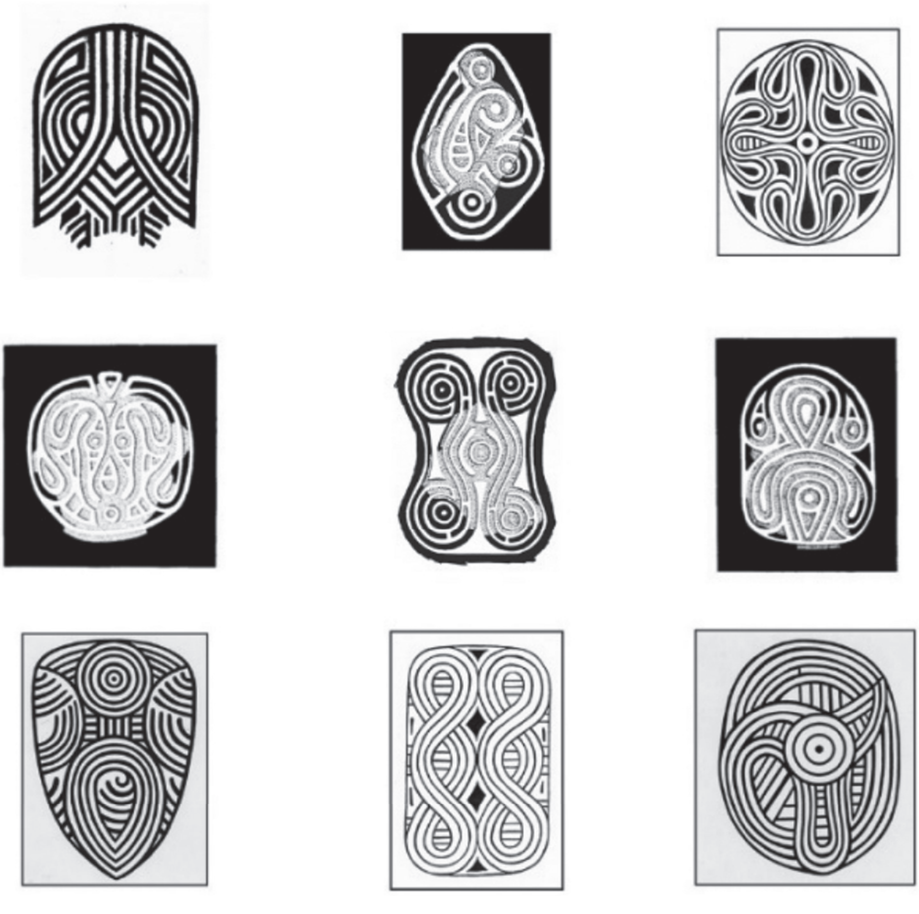


Figure 1. Paddle design reconstructions by Broyles (Row 1, design #2, BBFCL334; Row 2, all, BBP14-5, BBME26_4-1, BBP17-1), Snow (Row 1, design #3, FS007; Row 3, all, FS338, FS068, FS000), and an unknown llustrator (Row 1, design #1).

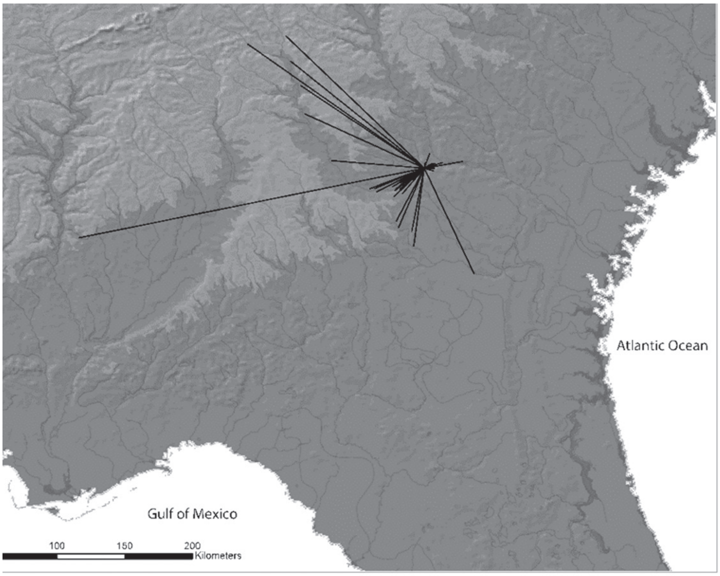


Figure 2. Design matches between Milamo (9WL1) and other sites in the region. Black lines represent one or more paddle design matches. Most matches with Milamo occur within the Ocmulgee-Big Bend (Stephenson, et al.).

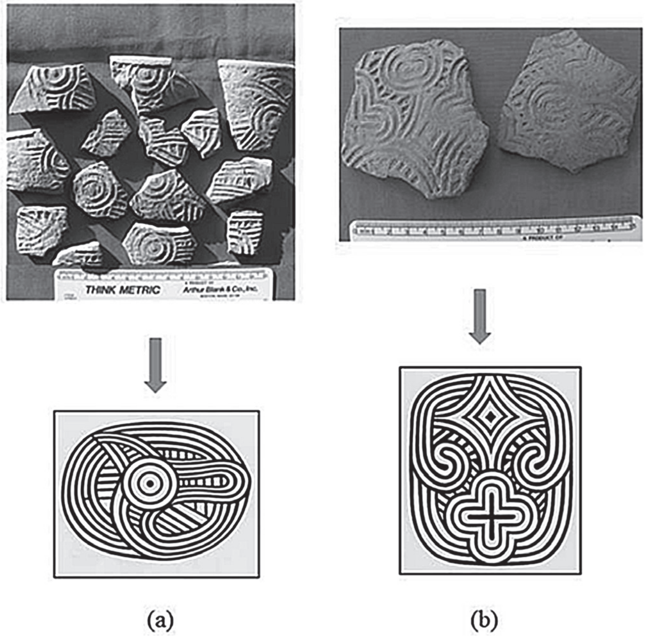


Figure 3. Sample pottery sherds (top) and their underlying wooden paddle designs (bottom). Two pottery sherds in (b) contain a composite pattern, resulting from the multiple applications of the carved paddle with partial spatial overlaps. Original designs reproduced with permission, courtesy of Frankie Snow, South Georgia State College.

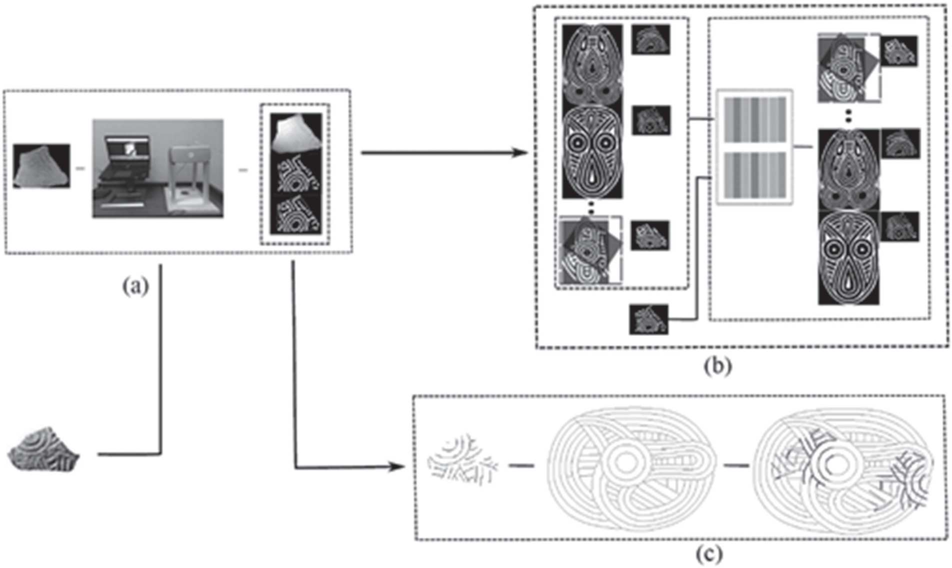


Figure 4. An illustration of a framework on identifying the underlying design for a sherd. (a) Extract a curve pattern from a sherd. (b) Identify the underlying design for a sherd with a non-composite pattern. (c) Identify the underlying design for a sherd with a composite pattern. Original design reproduced with permission, courtesy of Frankie Snow, South Georgia State College. Chamfer Matching algorithm (Barrow).

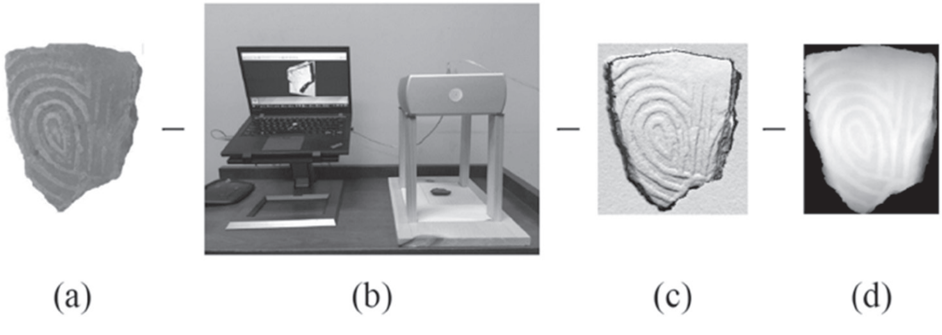


Figure 5: An illustration of scanning sherds for depth images. (a) RGB image of a sherd. (b) Setup of a 3D scanner. (c) 3D point cloud of the sherd surface obtained by the 3D scanner. (d) Depth image of the sherd surface: pixel intensity represents the depth value at a location.

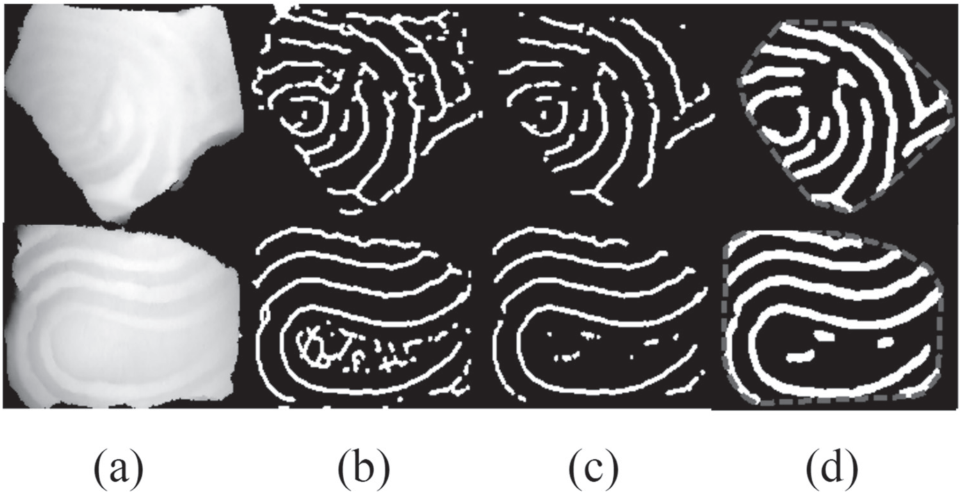


Figure 6. An illustration of segmenting curve patterns from sample sherds. (a) Depth images of sherds, where darker pixels have larger depths. (b) FCN-extracted curve skeletons. (c) Refined curve skeletons by using a dense prediction CNN. (d) Final segmented curve patterns with recovered curve width, masked by the sherd boundaries.

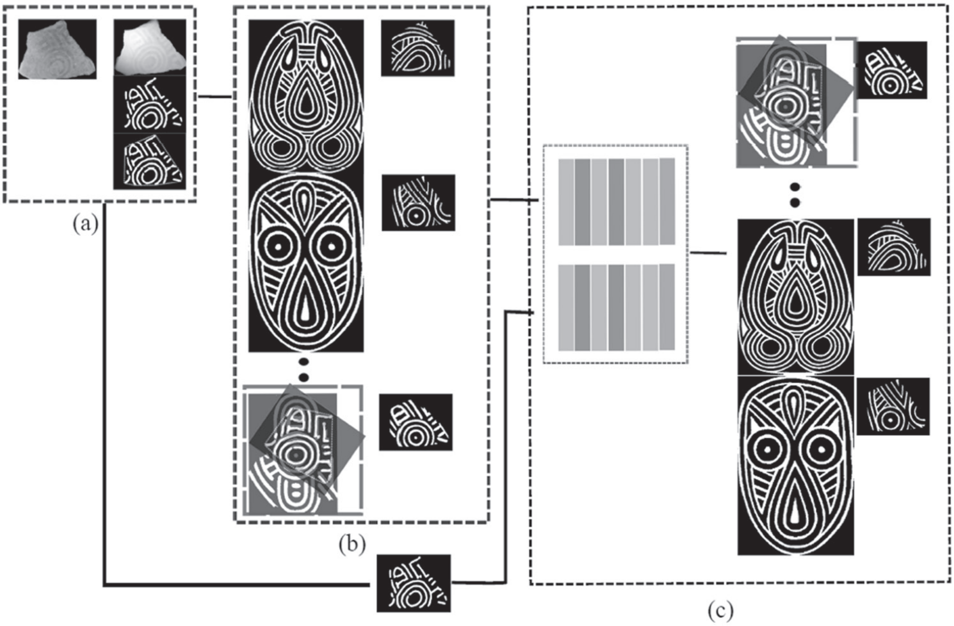


Figure 7. An illustration of the full pipeline of identifying design for a sherd with non-composite patterns. (a) Curve pattern segmentation from a sherd. (b) Stage 1: template matching with all the designs for selecting a small set of candidate matchings of the input sherd. (c) Stage 2: CNN-based re-ranking of the candidate matchings. Correctly matching design is shown in box, which is ranked low in Stage 1 but ranked at the top in Stage 2. Original design reproduced with permission, courtesy of Frankie Snow, South Georgia State College.

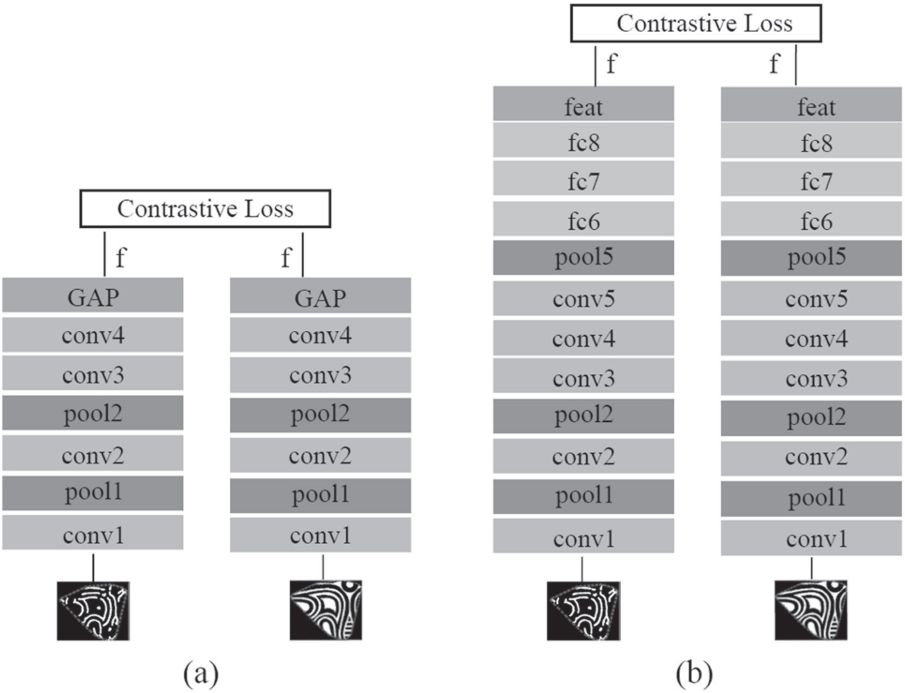


Figure 8. An illustration of the dual-source CNN architectures: (a) the proposed CNN and (b) AlexNet.

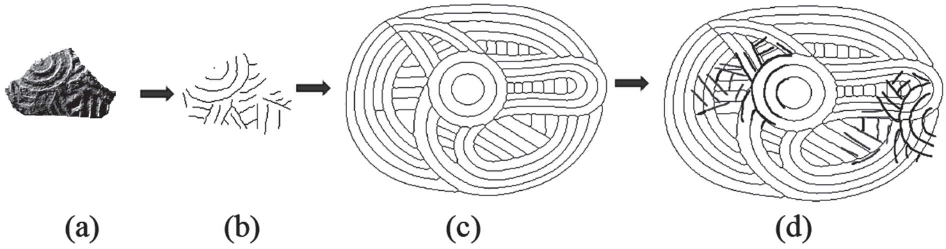


Figure 9. An illustration of identifying the underlying design for a sherd with a composite pattern. (a) A sherd image. (b) Curve extraction from a sherd. (c) Curve extraction from a design. (d) A sherd matching to two locations on a design. Original design reproduced with permission, courtesy of Frankie Snow, South Georgia State College.

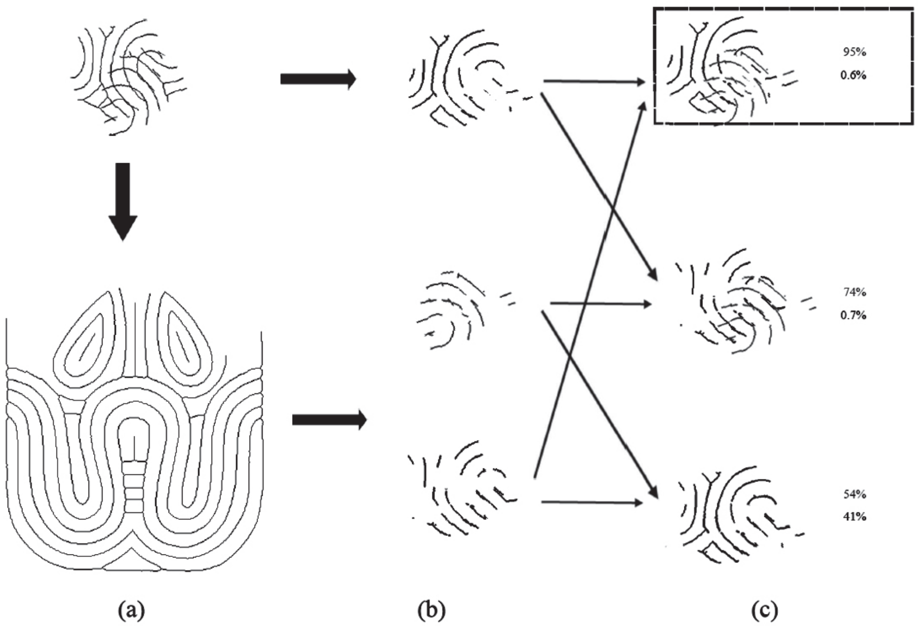
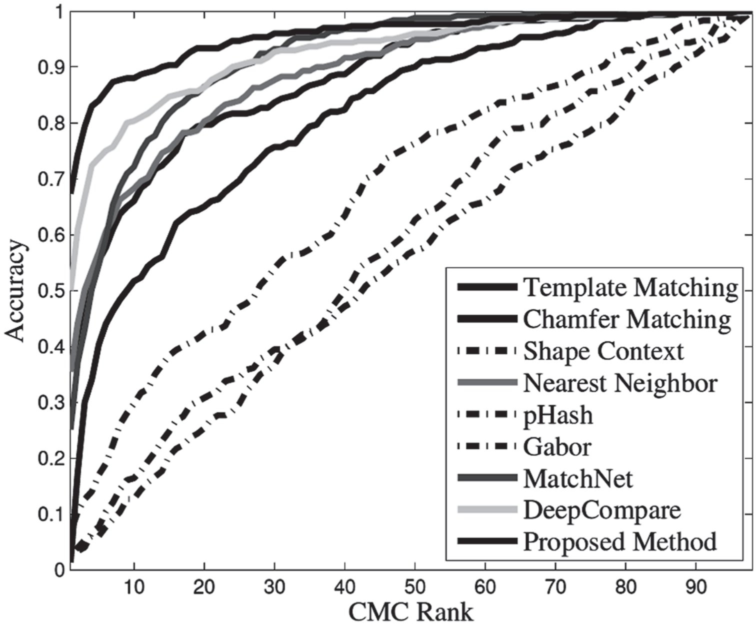


Figure 10. The process of combining candidate components for matching to a design. The optimal result is indicated in the box. (a) Matching a sherd pattern (top) to a design pattern (bottom). (b) Candidate Components. (c) Combining candidate components (best matching is shown in box). Original design reproduced with permission, courtesy of Frankie Snow, South Georgia State College.



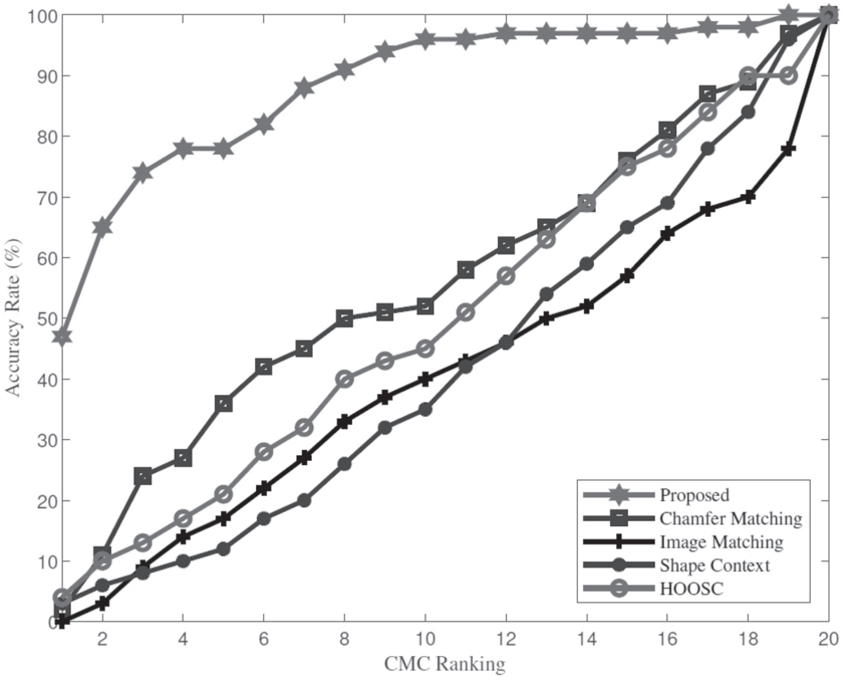


Figure 11. CMC curves of the proposed method and the comparison methods. Up: results on non-composite sherds. Down: results on composite sherds.