

Inferring the Lexicon of Ornament

Yuchen Yang¹, Melanie Gibson², Salim Arslan³, Arianna Salili-James⁴, Elena Cinelli⁵,
Theano Siarafera⁵, Imogen Brown⁵, George Manginis⁶, and Armand M. Leroi¹

¹Department of Life Sciences, Imperial College London, London SW7 2AZ, UK

²School of Arts, SOAS, London University, London WC1H 0XG, UK

³Department of Computing, Imperial College London, London SW7 2AZ, UK

⁴Department of Mathematics, Brunel University, Uxbridge, UB8 3PH, UK

⁵History of Design, Royal College of Art, London, SW7 2EP, UK

⁶Benaki Museum, Athens, 106 74, Greece

February 26, 2021

Abstract

Ornament has a language. Its lexicon are the motifs — decorative units — transmitted from artist to artist; its grammar how they are arranged on a given artefact. Here we develop statistical methods to study the immense diversity of ornament. We first use a deep convolutional neural network to isolate ornamental motifs at a large scale. We then analyse the colours, shapes and sizes of motifs and show how these features can be combined to infer the lexicon itself. We apply these methods to a famous body of ornamental objects, Iznik ceramics, many of which were made for the Ottoman court in Istanbul. We show that this lexicon evolved in the course of the 16th century and how, when combined with machine-learning techniques, it can be used to date Iznik artefacts. That is, we develop a computational connoisseur. Our methods can be applied to many kinds of ornament and open this vast domain of art to scientific study.

Introduction

We wish to understand the evolution of ornament: the origins of the beautiful patterns that fill our world, and the forces that produced them. Ornament is the 3-d ocelli and tiger stripes of the male Argus pheasant’s tail, the purple splash on a Song dynasty vase, and the gilt swirls of a Meissonnier soup tureen. But it’s not an actual tiger’s stripes, a M1 Abrams tank’s three-colour paint job or the pentagonal blob on a package of Persil for, while their primary purpose is to fool or attract the eye, ornament’s is to delight.

Ornamental patterns are typically composed of stereotyped sub-patterns or *motifs* [1, 2]. Some motifs, such as the “lazy wave” pattern on the rim of an Apulian skyphos, are simple geometries; others, such as the “the parasol ladies”, an orientalist image painted by Dutch artist Cornelius Pronk in the 1730s for the Far East porcelain trade, are complex figures. Motifs are sometimes repeated within a given artefact, arranged in hierarchies as in a Turkmen rug, symmetrically as in the ceiling of a Safavid mosque, or all three. These properties are not universal. Rococo ornament eschews symmetry and so do the marginal illuminations in *The Book of Kells*. A Chinese export plate has only one instance of its “parasol ladies”, but it is a motif nonetheless since the ladies, a pair of Japanese geishas, appear with little variation on thousands of them.

Motifs are designed to be reproduced. They are the fundamental units of transmission of ornamental art, the things that are copied by generation upon generation of artists and that, however they may be altered or combined with others, maintain their visual identity over time. If the way that they are arranged is an ornamental style’s grammar [3], then the motifs themselves are its lexicon. Scholars, tracing their etymology, have found that some are very ancient. The alternating flower-scroll borders found on modern Persian rugs are related to those carved on Diocletian’s 3rd century CE palace in Split, and to those painted on 7th century BCE Greek vases and, deeper yet, to the lotus-flower motif found on Dynastic Egyptian tombs [1, 4, 5]. But now we are talking more about a family of motifs linked by common ancestry. Even recent motifs have sometimes travelled far. Pronk’s “parasol ladies”, first used on export porcelain made in Jingdezhen and Deshima, was adopted by European artists who applied them to Delft faience, Venetian porcelain and, by 1800, English Spode [6].

As evolutionary biology begins with alpha taxonomy and comparative linguistics with lexica, so the scientific study of ornament begins with the identification and classification of motifs (cf. [7]). The scale of the task is daunting. Between 1600 and 1800, for example, the Dutch East India Company imported some 43 million pieces of Chinese and Japanese porcelain to Europe. Hundreds of thousands of them survive in public museums and private collections, and the diversity of their designs is vast [8]. Art historians can readily identify motifs by eye, but to enumerate and quantify the motifs of even a modest ornamental tradition would require an army of them. Lacking one, here we outline a computer-vision approach to do so.

We focus on an assemblage of ornamental artefacts famed for their beauty: Iznik ceramics (Figure 1A-H) [9–14]. Beginning in the 1480s artisans in Iznik, a town near Istanbul, began to produce lead-glazed fritware for the Ottoman court. Some probably worked in Istanbul itself [15]. Under the direction of Baba Nakkaş (fl. 1510-1520), a court designer, their first ceramics combined Chinese motifs and designs derived from silver, in a blue-and-white palette. A decade later they also made direct imitations of the blue-and-white Yuan and Ming porcelain prized by the Ottomans. Over time they added turquoise, red, emerald green and other colours to their palette and expanded their repertoire of motifs. Around 1535 Şahkulu (fl. 1526-1556), another court designer, introduced the Persian saz, an intricate, sinusoidal, leaf; his pupil and successor Kara Memi (fl. 1545-1566) developed a more realistic floral lexicon with tulips, carnations, roses, hyacinths and prunus blossom displacing the earlier Chinese cloud bands and fantastical floral designs. Besides tableware, Iznik artisans also produced ranges of tiles for the mosques, tombs and hammams built by the architect Sinan for the sultans and their courtiers. But then the Ottoman court began to feel the pinch,



Figure 1: Iznik and Iznik-style ceramics. **A.** tile, Iznik / Istanbul, 1520–1540; **B.** plate, Iznik, 1540–1560; **C.** plate, Iznik, 1550–1555; **D.** bowl, Iznik 1540–1560; **E.** plate, Iznik, 1570–1600; **F.** tile, Iznik, 1570–1575, **G.** tile, Iznik, c. 1580; **H.** tile, Iznik, 1600–1630; **I.** tile, Damascus, 1550–1600; **J.** plate, Florence (Cantagalli), c.1892; **K.** plate, Paris (Deck), c. 1870; **L.** tile, Derbyshire (Pilkington), c. 1862. See Table S1 for image sources.

commissions declined, and the potters turned to mass-production of tableware. In the 17th century production declined and, in 1719, the kilns at Iznik went cold.

The tiles, however, are still Istanbul’s glory. And Iznik style lived on (Figure 1I-L). In the mid-16th century Iznik artisans were sent to Damascus in Syria where they produced similar designs in a different palette [16]. Around the same time workshops in Florence and Venice were producing looser imitations [17, 18]. Between the early 16th and late 18th centuries a closely related tradition rose and fell in Kütahya in Western Turkey [19–21]. The rediscovery of Iznik by scholars and collectors in the late 19th century inspired a new wave of Iznik-derived ceramics [10, 22–28]. Ulisse Cantagalli in Florence, Theodore Deck in Paris, Armenian artisans in Kütahya and Athens, and the I.C.A.R.O. company in Rhodes all produced close imitations. In England William De Morgan borrowed Iznik motifs for his Arts-and-Crafts tiles, and in Holland the Arnhemsche Fayencefabriek gave them an Art Deco twist. There are Iznik-style tiles by Pilkington, an English firm, at the bottom of the North Atlantic on the walls of the hammam of the wreck of *RMS Titanic*.

The diversity of Iznik ornament is immense: Istanbul’s Rüstem Pasha mosque alone has thousands of tiles of at least 80 different designs [29]. A comprehensive study of the stylistic diversity of Iznik ceramics requires, then, automated methods. Before sketching our approach, we define some terms. An “artefact” is a whole plate, jug or tile or, sometimes, a few tiles. Each artefact is decorated with many “ornamental objects” or simply “objects”, for example, individual tulips, and each object can be assigned to a “motif”, a group of objects whose shape, colour and size resemble each other. Each motif, in turn, belongs to a motif-family such as “tulips”. Here, then, we begin by developing a deep learning model capable of automatically identifying and segmenting ornamental objects belonging to three families of motifs: tulips, saz-like leaves and carnations. We then show how these objects can be classified into distinct motifs based on their colours, relative size and shape. Finally, we investigate whether a machine-learning model trained on the relative abundance of motifs can be

used to classify artefacts into their approximate dates of production. In doing so we identify part of the lexicon of Iznik ornament and show how it evolves. Collectively these methods are a step towards the large-scale, quantitative analysis of the evolution of ornament.

Results

Using deep learning to identify ornamental objects

We began by investigating the ability of a deep learning model to automatically segment ornamental objects. To do this we built a corpus of 429 images from museum digital catalogues and monograph illustrations, each of which shows a single, unique, artefact. We then visually identified all tulips ($N = 1991$), carnations ($N = 461$), and saz-like leaves ($N = 1642$) and manually segmented them (see S.I. section S2 for definitions of motif families). We based our segmentation models on Mask R-CNN, a deep learning framework widely used for segmentation tasks [30–36]. The Mask R-CNN was pre-trained on a standard segmentation dataset, MSCOCO 2017, provided by the Detectron2 API, and the models were then fine-tuned to find all three of our motif-families — tulips, carnations and saz-like leaves — on our artefact images. To do this we trained each model on 90% of the images, chosen at random, predicted the objects present in the remaining 10%, and replicated this procedure 10 times.

Each prediction locates a putative object in an image with a mask, identifies motif-family to which it belongs, and has a prediction score that tells us how good the model thinks it is. Figure 2A shows four tulips on a plate from the Ashmolean successfully predicted by one model. To evaluate each model we calculated the average precision, AP of our predictions. To calculate AP it is necessary to specify a criterion by which a given prediction is deemed as being as acceptable. The standard way to do this is to set a threshold for the intersect over the union, or IoU , a metric that varies between 0 and 1 and gives the extent to which each prediction overlaps with a ground truth object. As the IoU threshold increases AP declines since fewer predictions are accepted, that is, are judged to be true positives. A table in Figure 2A reports AP s for two conventional thresholds: $IoU = 0.5$, and for all $0.5 \geq IoU \leq 0.95$. The overall mean performance of the model has an $AP_{0.5-0.95}$ of around 60% and is comparable to that of object detection tasks with similar challenges [32, 36]. But we found that the models were better at segmenting some motif families than others. While it was able to segment tulips and carnations with relatively high precision it wasn't as good at saz-like leaves. This is not due to a lack of saz-like leaves in our training set, but rather their heterogeneity, a point which we elaborate below.

How useful is our deep learning model? Suppose we obtain an entirely new set of 1000 artefact images, say, tiles from the Rüstem Pasha mosque. If the distribution of motifs on these new images is the same as that of our test set then they will collectively have about 7867 tulips, carnations and saz-like leaves on them. Applying any one of our models we will find that it makes many more predictions than there are objects in the images and that most of these predictions are poor: they are false positives in that they either predict an object where none exists or predict its outline only very roughly. We remove most, but not all, poor predictions by accepting only those that exceed some threshold on the prediction score, s . For example, taking one of our models, Mask R-CNN-Iznik.1, if we set a filter at $s \geq 0.075$, we will have 7617 predictions for 5634 or 72% of the objects. Those predictions will, however, include some which the model thinks are good but are in fact bad. To eliminate bad predictions we visually inspect them keeping only those in which the prediction and the predicted object belong to the same motif-family (e.g., are both tulips) and overlap by about 75%. Following this procedure we estimate that we will obtain good predictions for 4234 or 54% of the 7867 present in the original images with a bias in favour of tulips and carnations and against saz-like leaves. Figure 2B shows the predictions and ground truth for the 255 objects in the test set that passed these filters.

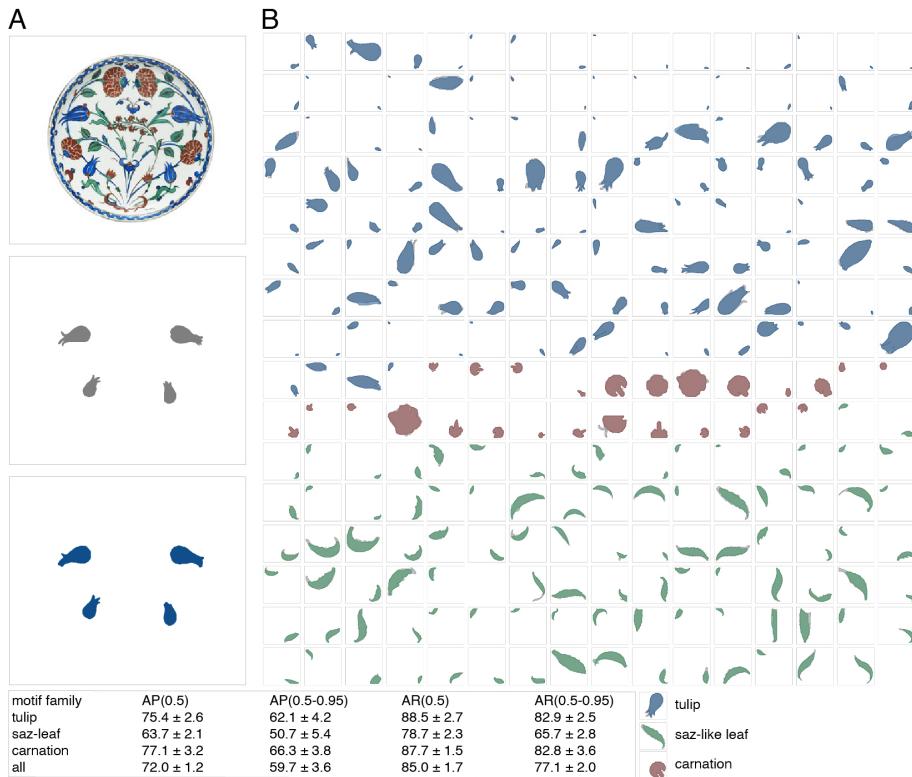


Figure 2: Segmenting ornamental objects by deep learning. **A.** A successful example. Top panel. a plate from the Ashmolean (AM05/EA1978.142) containing four tulips. Middle panel. Manually isolated ground truth objects. Bottom panel. Objects predicted with score $s > 0.9$. **B.** Successfully segmented objects from one of the three test sets. Predictions filtered by prediction score $s > 0.1$ and $IoU \geq 0.75$; ground truth in grey overlain by coloured prediction.

This procedure requires visually comparing thousands of predictions to the original images. Tedious — but still much easier than segmenting the objects by hand. But what of the 46% of objects that were either predicted poorly or not at all? If they are perfectly good objects, perhaps of a particular kind, then we need a better model; if, however, they are fragments or the occasional wonky object (for not every Iznik artist was a master all the time) then it is fine. To find out we examined the objects that failed our filters in Mask R-CNN-Iznik.1. For tulips and carnations we found that failures were nearly always closely related to successfully segmented objects. For example, VA15, a plate from the Victoria & Albert Museum, has 20 micro-tulips of which the model failed to segment 9. For saz-like leaves we found that the model sometimes failed to accurately segment complex arrangement of leaves such as in a rosette (cf. Figure 1E) or else when serially concatenated to make a larger leaf, but that it also found other instances of these. The model also sometimes failed to find partial objects of all families. These results suggest that some improvement might be obtained by re-defining the objects or expanding the training set (particularly of saz-like leaves) but also that our model, and this procedure, can identify nearly all motifs, but not all instances of them, in a large set of new images.

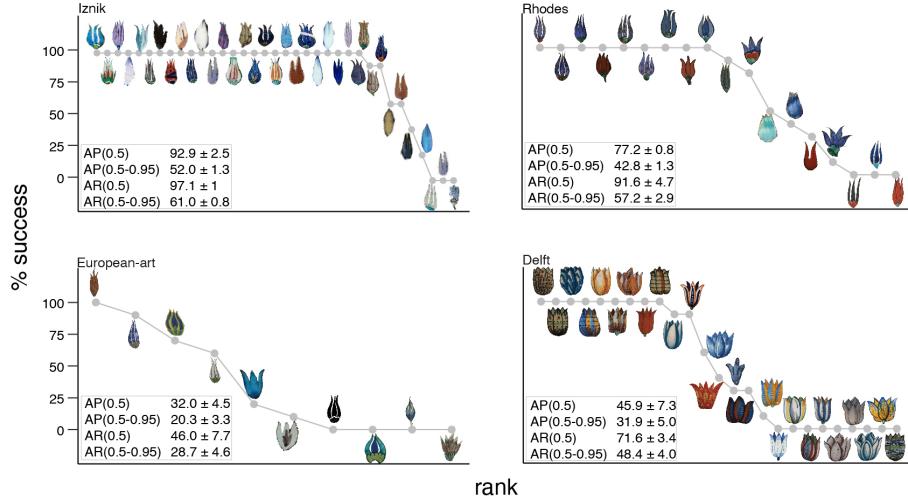


Figure 3: Performance of Iznik-trained deep-learning models when finding selected Iznik and non-Iznik ornaments. Success rates across 10 replicate models for each tulip. The tulips themselves are shown scaled to unit size. In the tables AR and AP refer to Average Precision and Average Recall respectively and the two levels to $IoU \geq 0.5$ and $0.5 \leq IoU \leq 0.95$

Of course, new images may contain new motifs — entirely new kinds of tulips, say — and to find them we might need to expand our training set and retrain on them. To get a sense of how widely applicable our current models are, we asked whether they could recognize motifs in other, related, stylistic traditions or even in unrelated traditions that have independently adopted similar motifs. We examined two descendant styles. The first consists of ceramics produced between 1930 and 1988 by the ICAROS factory on the Greek island of Rhodes whose ornamental schemes were closely based on the 16thC. Iznik ceramics abundant in the houses of the local bourgeoisie. The second, looser, group of ceramics were produced between the 1870s and 1930s by factories in England and the Netherlands and were plainly inspired by Iznik insofar that they have many of its characteristic motif-families — tulips, saz-like leaves and carnations — but modified by contemporary aesthetic trends such as arts-and-crafts and art-deco. We tested the ability of our models to identify the tulips in these “Rhodes” and “European-art” traditions. We chose tulips because our model is quite good at finding them in Iznik, but also because we wanted to study a third ornamental tradition in which they are found: 17th century Dutch Delft. The Delft and Iznik ceramic traditions, near-contemporaries, both consist of tiles as well as various vessels made of lead-glazed fritware often

decorated with flowers. Yet despite this — and despite the abundant trade between Amsterdam and Istanbul — scholars believe that the Delft ornament is only remotely related to Iznik, and that its tulips came from contemporary Dutch floral still-life paintings, illustrations in plant books and growers’ catalogues rather than Turkish tiles [37]. Delft tulips are a home-grown product of the “tulipmania” that swept Golden Age Holland.

Given these relationships we expected that our models would be good at identifying Rhodes tulips, do less well at identifying their more distant European-art relations, and be poor at identifying unrelated Delft tulips. Since we had relatively few artefacts from these traditions, we curated and edited the images so that each had only one, whole, tulip clearly different in form or colour from all others. These datasets consist of 18 Rhodes, 10 European-art and 25 Delft images and were compared to a dataset of 28 Iznik images (two per motif; see below) treated in the same way (Figure 3). Considering average precision, $AP_{05-0.95}$, and average recall, $AR_{05-0.95}$ our models are capable of discovering, in decreasing order, Iznik, Rhodes, Delft and European Art tulips (Figure 3 inset tables). Looking at these results another way, where $91 \pm 9\%$ of Iznik tulips were found ($IoU \geq 0.75$) by at least one replicate model, the equivalent percentages for Rhodes, Delft and European-art tulips were, respectively, 83 ± 17 , 64 ± 19 and $60 \pm 30\%$. Thus, as expected, Rhodes tulips are most similar to Iznik but, contrary to our expectations, Iznik tulips are less similar to their European-art descendants than they are to unrelated Delft tulips. Comparing non-Iznik tulips that were consistently found to those that were not suggests that the model is particularly poor at discovering the more naturalistic Delft tulips (Figure 3D).

Discovering the lexicon of Iznik ornament

The tulips, carnations and saz-like leaves depicted on Iznik artefacts come in a bewildering variety of colours, shapes and sizes. To classify them into motifs we estimated their colours, relative sizes and shapes and then applied an unsupervised clustering method to these data. In this part of our study we used a subset of the artefacts and objects in our corpus requiring that the images of the artefacts be of high quality and that the objects themselves be complete, that is, not truncated by continuing on an adjacent tile or around the back of a jug. In all, we classified 2357 objects drawn from 268 Iznik artefacts: 982 tulips, 255 carnations, and 1120 saz-like leaves.

We began by studying their colours. Between 1480 and 1700 the colours used by Iznik artisans expanded from a few shades of dark cobalt-blues to a polychrome palette of about a dozen colours, though not all were used at any given period as colours fell in and out of fashion (S.I. Table S3). In addition to these colours the base (slip) colour of Iznik ceramics, although typically an off-white, varies. Although scholars generally speak with confidence of the colours of Iznik artefacts, in fact any given colour, say “cobalt-blue”, comes in many different shades depending on the exact composition of the pigment, how it was applied, and the firing conditions of the artefact. In addition, our artefacts were photographed in many places and times without colour standards giving rise to another source of variation.

This variation makes it hard to objectively say whether any two artifacts have the same colours or not. We approached this problem by mapping the wide range of pixel colour values present in our images to a standardized palette. We first tried to construct such a palette by simple unsupervised clustering on colour values. This approach, however, proved unsatisfactory for it failed to isolate several distinctive, important, but uncommon, colours (see S.I. “Unsupervised palette discovery”). For this reason we adopted another approach in which we mapped pixels to a palette defined *a priori* by an expert.

To construct this “supervised” palette we picked out 73 images that, together, represented the full range of Iznik ware. We then asked an expert (MG) to identify the colours she thought were present in those images and where they occurred. She identified 11 distinct colours. (Our unsupervised analysis suggested a similar number, but divided the colour space differently.) We then sampled

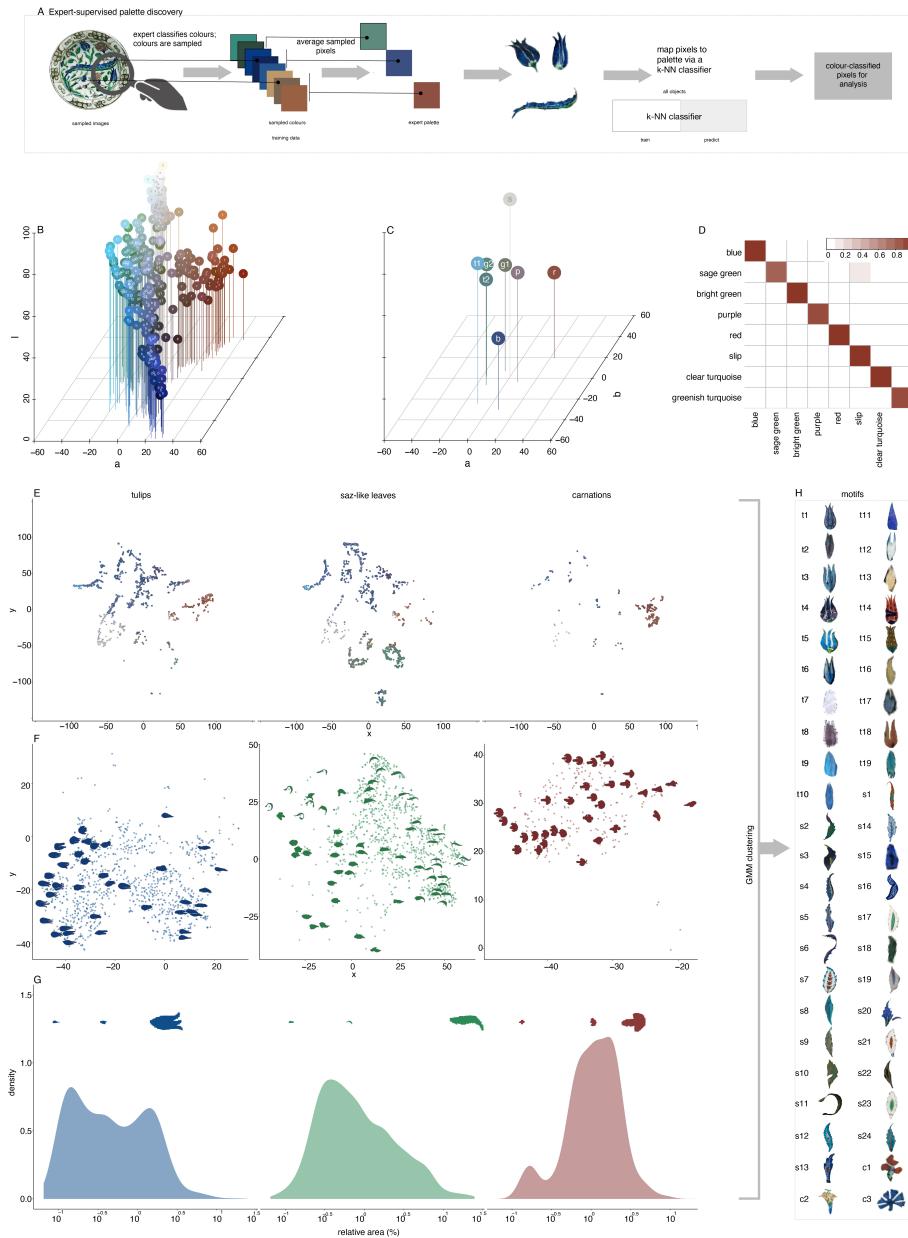


Figure 4: Phenotyping ornamental objects and classifying them into motifs. **A.** Flow chart showing the discovery of the expert-supervised palette of ornamental objects. **B.** Expert-supervised palette. Each point shows the median value in CIELAB space of the pixels sampled for each colour in the 73 individual artefacts used to establish the supervised palette: b: blue, g1: sage green, g2: bright green, p: purple, r: red s: slip, t1: clear turquoise, t2: greenish turquoise. **C.** Expert palette: median values of the training set colours over all artefacts, labelled as in A. **D.** Confusion matrix of the k-NN classifier used to assign the expert palette to individual objects. The k-NN had an overall accuracy of 96% once pixels assigned to colours with a low probability were excluded. **E–G** Distributions of object phenotypes. From left to right: tulips, saz-like leaves and carnations. **E.** Colours. The colours of each object were inferred from the expert-supervised palette; points show the two most common colours of object positioned in a common colour space reduced to two dimensions by tSNE. **F.** Shapes. Points show the positions of individual objects are in a common colour space reduced to two dimensions by tSNE from Currents Distances. A random selection of object outlines is plotted for reference. **G.** Relative size. Density distributions (log10 scale) of objects. Scaled object outlines from the extremes and medians of each distribution are shown for reference. **G.** Motifs discovered by GMM clustering on phenotypes. 46 motifs were discovered in all: 19 tulips (t), 24 saz-like leaves (s) and 3 carnations (c). The highest probability object of each cluster is shown, scaled to unit length.

about $2.7 \cdot 10^6$ pixels in the regions she indicated and, from them, obtained an average (median) set of CIELAB values for each colour to define the palette. Having assigned each pixel to one of these 11 colours we trained a k-NN classifier on 70% of them, reserving the rest of the test set. At this point we found that we had to modify our expert’s palette in two respects. First, she identified a black that was sometimes used to outline individual objects. This pigment was, however, used so sparingly that we could neither convincingly sample it nor distinguish it from very dark greens or blues that were also sometimes used in outline, so we excluded it. Second, she identified three shades of blue: dark, cobalt and pale, but we found that we could not train our colour-classification model to distinguish between them with any accuracy, so we combined them into a single blue. The resulting palette has eight colours that we call: blue, sage-green, bright-green, purple, red, clear-turquoise, greenish-turquoise, and an off-white slip colour (Figures 4B, C). Using this modified palette we found that the resulting model had, once the blues were combined, an overall accuracy of 86%; excluding pixels with a classification probability < 0.75 increased the overall accuracy to 96% (Figure 4D). Having obtained an accurate model, we then used it to predict the palette colours, and then their frequencies, present in all 2375 ornamental objects in our data set; Figure 4E shows the distribution of the objects in a common two-dimensional colour space as inferred by this approach. It also shows how the Iznik artisans used colours in the different families of motifs. For example, there are many tulips and carnations that are mostly red but relatively few saz-leaves; and, while there are many tulips and saz-like leaves that are mostly clear-turquoise, very few carnations in our data set are.

To describe the shapes of our tulips, carnations and saz-like leaves we automatically extracted their contours, standardized their size and orientation and contour startpoints, and then estimated their distance using a Currents algorithm [38] which yielded 18 principle components that described 99% of the variation in non-Euclidean shape space (Figure 4F). Then, as a measure of size, we obtained the area of each object relative to the artefact that bears it (Figure 4G). Finally, we combined the colour, shape and size data and clustered the objects using a Gaussian Mixture Model and found that, 19 tulip, 23 saz-like leaves and 4 carnation clusters were optimal (Figure 4H). Each of these clusters, then, is a motif; collectively, they are an ontology of ornamental motifs and, as such, part of the lexicon of Iznik ornament.

Lexical evolution

We do not claim that our motif ontology is definitive. We measured and clustered our ornamental objects using certain methods; other methods will yield other classifications. However, a classification’s merit lies in its utility, and we would like to use our motifs to classify the artefacts upon which they are found. In particular, we would like to use our motifs to date artefacts automatically. To find out whether they can be used to do so we asked an expert (MG) to classify our artefacts by date of manufacture. She used the criteria of traditional connoisseurship — colour scheme, geometrical relationships of the motifs, artefact shapes, and the presence of various signature motifs — to produce two groups of artefacts: “mid-16thC” (1520–1560; $N = 55$) and “late-16thC” (1561–1600; $N = 179$). (She also identified a third group made before 1520, but there were so few of these that we did not consider them further.) Even though our expert did not know our motif-ontology, we found that 19/27 of our motifs showed statistically significant ($\alpha = 0.05/2$) differences in frequency among the date-classes (Figure 5A). We then tested the ability of a supervised classifier to classify artefacts into date-classes based on the frequencies of their motifs. Since we have many more late-than mid-16thC artefacts, we generated a thousand date-class balanced data sets of 110 artefacts each and applied a k-NN classifier to each using a train:test ratio of 0.8:0.2, so that each artefact was tested against at least one hundred different training sets. We found that our models had an average precision (accuracy) of $AP_{mid} = 0.75 \pm 0.01$ (95% CI), $AP_{late} = 0.85 \pm 0.01$ and a combined precision of $AP_{combined} = 0.80$. The comparable recall values were $AR_{mid} = 0.85 \pm 0.01$, $AR_{late} = 0.79 \pm 0.01$

and $AR_{combined} = 0.82$. Figure 5B shows a selection of the most successfully classified artefacts in each date-class. When dating the artefacts our expert used much more information than our k-NN models did, and our training set was modest in size; nevertheless these results show that our motif ontology captures at least some of the evolving lexical structure of Iznik ornament and that it can be used to classify the artefacts that bear them.

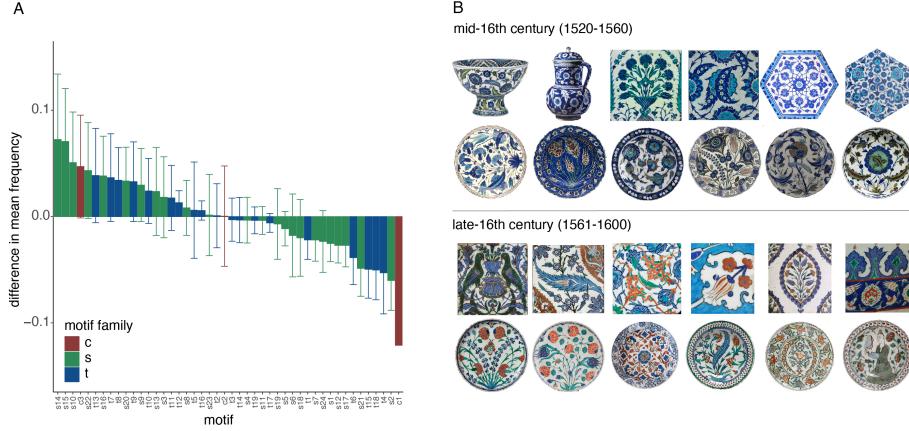


Figure 5: Classification of artefacts by date. **A.** Difference in frequency (mid–late) of motif frequencies between mid- and late-16thC artefacts: means and 95% confidence intervals ($N_{mid} = 55$; $N_{late} = 179$). **B.** A selection of artefacts that were correctly classified in at least 75% of the replicate k-NN runs in which they appear.

Discussion

The development of image analysis tools, and the availability of large corpora of digital images, has enabled a new, scientific, art history that seeks to describe patterns of artistic evolution and their causes. Most studies of this sort have described the evolution of Western figurative art using high-level stylistic features such colours, texture or visual complexity [39–45]. The study of ornamental style, however, requires a different approach.

The fundamental units of ornament are motifs: they are the things that are copied, generation upon generation, by artists and which maintain their visual integrity over hundreds, or even thousands, of years. For this reason we developed ways to automatically isolate ornamental objects and classify them into motifs. Having done so we showed that the motifs found in our sample of Iznik artefacts changed in frequency over the course of the 16th century. That implies that the distribution of motifs on a given artefact — a plate, say — reveals when it was made. Using a simple machine-learning classifier, we were able to date artefacts automatically with some accuracy ($AP_{combined} = 0.80$) even though we studied only three of the many motif families that the Iznik artisans used.

As such, our study opens the door to a computational connoisseurship of ornamental style. Few Iznik artefacts bear dates. Scholars therefore date them by comparing their stylistic properties with the few objects that can be securely dated, such as a mosque lamp, now in the British Museum, that has the year 956 AH — 1549 CE — painted on it. The date intervals that we used were inferred in just this way which is why they span three or four decades each. However, many thousands of tiles remain attached to the mosques and hammams that the Ottomans built and, thanks to the archives that their bureaucrats left behind, we know when they were built and hence — with some caveats, for tiles were occasionally reused — when the tiles that adorn them were made. An analysis of these, rather than the tableware that dominates our sample, should allow us to construct a much more detailed stylistic chronology than has been possible until now, one that can be used to give

probabilistic estimates of when the many isolated artefacts found in collections throughout the world were made. Until now, dating Iznik, like any unsigned art, has been the province of experts. While we doubt that the connoisseur’s keen eye can be replaced, we do think that statistical approaches to art history based on publicly available datasets, rather than hermetic knowledge, will soon prove their worth [46].

Our interest, however, lies less in making computational connoisseurs than in charting the deep history of ornament. The production of ceramics at Iznik was but a brief episode in ornament’s long history. Scholars, pointing to the vast collection of Ming porcelain in the Topkapi Palace or the Persian origins of some of the Ottoman court designers, have shown that Iznik combines motifs drawn from Chinese, Persian as well as other sources [9–11, 26, 47]. And, long after tulips were no longer painted at Iznik, the style spread from Northern Europe to Central Asia, appearing sometimes as frank imitations and sometimes as vague allusions to a long-forgotten source [16–18, 28]. It is this transmission of motifs that fascinates us, and we envision building a comprehensive library of training sets and machine-learning models based on many ornamental traditions to track their spread. Taking Owen Jones’ idea of ornamental styles as languages literally [1, 3], we suggest that, given suitable data, it should be possible to reconstruct their historical relationships using modern phylogenetic methods rather as comparative linguists have reconstructed the histories of the languages we speak (e.g., [48, 49]). In this view, motif-families are cognates or homologues. As such, their presence in a ornamental style marks its affinities rather as certain words mark the languages we speak.

In this spirit, Tehrani and colleagues have used phylogenetic methods to reconstruct the relationships of the ornamental dialects visible Iranian and Turkmen tribal textiles [7, 50–52]. However, where they scored their motifs by eye, we have shown how this might be done automatically at larger scale. Different methods of inferring genealogical relationships will be needed as well. The language of tribal textiles is largely transmitted vertically from mother to daughter. This means that their genealogy may well resemble a bifurcating tree which can be estimated using methods borrowed directly from evolutionary biology. The Iznik artisans, by contrast, worked in industrial-scale workshops directed by cosmopolitan designers who took their ornamental vocabularies from where they could, and many of their European successors did too. Thus a genealogy of Iznik and its kindred styles will certainly not resemble a tree but rather a network, one that spans many cultures, centuries and continents. Any influence-network construction method should allow for the possibility that motifs may be independently invented — as, apparently, tulips-on-tiles were invented in Turkey, and then again in Holland a century later.

Finally, we note that, while we have identified part of the lexicon of Iznik, we have not studied its grammar. By the “grammar of ornament” Owen Jones meant the “general principles in the arrangement of form and colour” by which ornamental motifs are combined to make a beautiful whole and that he, and later [1], explained in terms of visual psychology. Their evidence was, in the way of art historians, a multitude of suggestive examples, but it may be possible to infer these principles statistically by studying the geometrical arrangements of motifs on many artefacts. Whatever their nature they are surely part of the formula that has given these works of art, made in a provincial Asia Minor town more than four centuries ago, their enduring appeal.

Materials & Methods

Segmentation using a DCNN

Data

We obtained images of 429 Iznik artefacts from online museum catalogs and published monographs, each of which shows a single, unique, artefact. We then visually identified all tulips ($N = 1991$),

carnations ($N = 461$), and saz-like leaves ($N = 1642$) present in these images using the definitions of these motif families given in S.I. section [S2](#) and manually segmented them. Before submitting these images to the Mask R-CNN we split some of the larger images into sections, so that we ultimately used 575 images. The metadata and the segmented images are available at the following [Github](#) repository: [Armand1/The-Lexicon-of-Iznik-Ornament](#).

Model configuration

Our segmentation model is based on the Mask R-CNN [30] implemented on [Detectron2](#) and [PyTorch](#). We used a pre-trained resnet-50 and Feature Pyramid Network backbone as the baseline as it obtains the best speed and accuracy tradeoff across all available weights. Models were trained on the MSCOCO 2017 train/val dataset provided by the [Detectron2 API](#). The final base learning rate was set to 0.02 with an optimal performance, according to the test of learning rate from 0.05 to 0.0001 and the other training parameters were chosen to match with the hardware (Table [S2](#)). We ran this model with 10 fold replication, randomly allocating the images between training and test set each time. Model training and segment inference used [Python 3.6](#), [Torch 1.5](#) and [Detectron2 0.1.3](#) running on a single NVIDIA Tesla P100 GPU.

Model evaluation

To evaluate our model we calculated Average Precision, AP and Average Recall, AR . To do this we ranked the predictions by their prediction scores. We then set an intersection over union, or IoU , threshold at which to estimate AP or AR . IoU is a metric of the degree to which a prediction overlaps with its corresponding ground truth, and varies between 0 (no overlap) and 1 (complete overlap):

$$IoU = \frac{M_p \cap M_g}{M_p \cup M_g} \quad (1)$$

where M_p and M_g are the set of pixels that occur in the predicted and ground truth masks respectively. Given some IoU threshold, we judge each prediction to be either a true or false positive and then calculate the precision and recall rates over all ranked predictions as:

$$Precision = \frac{N_T}{N_P} \quad (2)$$

$$Recall = \frac{N_T}{N_G} \quad (3)$$

where N_T , N_P , N_G are the counts of true positives, predictions and ground truth for any IoU threshold and rank in our list of predictions. These values are then used to construct a Precision-Recall curve where AP , constitutes the area under this curve. We calculated AP and AR for IoU thresholds of 0.5 and 0.5–0.95 using 0.05 size bins and refer respectively to them as $AP_{0.5}$, $AP_{0.5-0.95}$, $AR_{0.5}$ and $AR_{0.5-0.95}$. To calculate these metrics we used the [COCO API](#).

Motif classification

Data

The objects — tulips, carnations and saz-like leaves — that we classified into motifs form a high-quality subset of all of the motifs in our data set. This is because, in order to find out how many different motifs we have, we only wanted objects that were complete (that is, not fragmented or partially occluded) and for which we had good photographs (that is, were not unduly dark or oddly coloured). In all, we classified 2357 objects drawn from 268 Iznik artefacts: 982 tulips, 255 carnations, and 1120 saz-like leaves.

Colour

To establish the supervised palette we asked an expert (MG) to identify the colours she thought were present in our corpus of images. She identified eleven colours in images of 73 artefacts. We sampled $2.7 \cdot 10^6$ pixels in the regions she indicated and, from them, obtained an average (median) set of CIELAB values for each colour to define our supervised palette. When making this palette, however, we ignored our expert in two respects. First, she identified a black that was sometimes used to outline individual objects. This pigment was, however, used so sparingly that we could neither convincingly sample it nor distinguish it from very dark greens or blues that were also sometimes used in outline, so we excluded it. Second, she identified three shades of blue: dark, cobalt and pale, but we found that we could not train our colour-classification model to distinguish between them with any accuracy, so we combined them into a single blue. The resulting palette has eight colours that we call: blue, sage-green, bright-green, purple, red, clear-turquoise, greenish-turquoise, and an off-white slip colour (Figures 4 B,C). Having labelled each pixel in this sample with a colour, we then used it to train a k-NN classifier, with $k = 100$ implemented in the `class` package in R [53] and, having trained it, used it to predict the colours of all pixels of all objects.

Shape

Each object in the motif-classification set was manually segmented and placed against a white background. We extracted their outlines using the marching squares method as implemented in the `Python sci-kit image` library. The resulting outlines were interpolated to give 240 points per outline, size-standardized, lightly smoothed, and roughly aligned within each motif-family using functions in the `Momocs` library in R [54]. We then undertook a further round of alignment using a custom function that: (i) rotated each object outline in a given family sequentially through 180° using the `rotate` function in `Momocs`; (ii) for each image rotation rotated the start point of its coordinates through all 240 possible positions; (iv) sought, for each object, the image and start point rotation whose y coordinates were most highly correlated those of a target object of the same motif-family. The result is a collection of outlines, one per object, with the same orientation and whose coordinates begin at the same landmark. We then obtained pairwise distances between all objects using Geometric Currents, a Diffeomorphic distance method implemented in `Python` [38]. To plot the positions of the objects in shape space we reduced the Principal Components produced by Geometric Currents to two dimensions using t-SNE implemented in the `Rtsne` package in R [55].

Size

The relative size of each object was obtained by dividing its area in pixels by the area of the artefact it was found on, for example, a tile or plate. In the case of tiles, which often come in panels, we estimated object size relative to a single tile. T

Clustering to find motifs

To obtain the motifs we combined the colour, shape and size data for all objects, log-transformed the relative size data and, for each motif-family, reduced its dimensionality by Principal Components Analysis using `princomp` in base R, keeping those PCs that accounted for 95% of the overall variance. We then clustered the objects within each family using a Gaussian mixture model (GMM) implemented in the R package `mclust`, finding the optimal clustering solution by BIC [56].

Artefact classification

We asked an expert, MG, to date our artefacts using the criteria of traditional connoisseurship, that is, without knowing our motif classification. When then estimated the frequencies of each motif in each artefact and used these data, and the date-classes, to train a k-NN model implemented in the `class` package in R [53] with $k = 12$, reserving 20% of the sample for the test set. We repeated this 1000 times, randomly allocating artefacts to the training and test set each time so that each artefact was tested against at least 100 different training sets. See [ref: an online depository] for origins on the images.

Acknowledgements

We thank Sarah Teasley, Chiara Barbieri and Dorothy Armstrong for helping to organize the Royal College of Art’s participation; Yanling Li for assistance in data acquisition; the Benaki Museum for supporting a workshop at the Leigh–Fermor House; and Yiannos Ioannidis for providing us with unpublished images.

References

1. Gombrich, E. H. *The sense of order: a study in the psychology of decorative art* (Phaidon, 1994).
2. Trilling, J. *The language of ornament* (Thames & Hudson, 2001).
3. Jones, O. *The grammar of ornament* (Day and Son, 1856).
4. Goodyear, W. H. *The grammar of the lotuses* (Sampson, Low, Marston, 1891).
5. Riegl, A. *Problems of style: foundations for a history of ornament* (Princeton University Press, 1992).
6. Jacob-Hanson, C. Spotted in England: Pronk's Chinese Parasol Ladies. *Northern Ceramic Society Newsletter*, 27–36 (Dec. 2018).
7. Tehrani, J. J. & Collard, M. Investigating cultural evolution through biological phylogenetic analyses of Turkmen textiles. *Journal of Anthropological Archaeology* **21**, 443–463 (2002).
8. Campus, J. v. & Eliëns, T. *Chinese and Japanese porcelain for the Dutch Golden Age* (Waanders, 2014).
9. Lane, A. *Later Islamic pottery. Persia, Syria, Egypt, Turkey* (Faber & Faber, 1957).
10. Atasoy, N. & Raby, J. *Iznik: pottery of Ottoman Turkey* (Laurence King, 1994).
11. Denny, W. *Iznik: The artistry of Ottoman ceramics* (Thames & Hudson, 2004).
12. Carswell, J. *Iznik pottery* (British Museum, 2007).
13. Ribeiro, M. Q. *Iznik pottery and tiles in the Calouste Gulbenkian Collection* (Calouste Gulbenkian Foundation, 2009).
14. Bilgi, H. *Iznik: the Ömer Koç collection* (Vehbi Koç Foundation, 2015).
15. Necipoğlu, G. From International Timurid to Ottoman: a change of taste in sixteenth-century ceramic tiles. *Muqarnas* **7**, 136–170 (1990).
16. Millner, A. *Damascus tiles: Mamluk and Ottoman architectural ceramics from Syria* (Prestel, 2015).
17. Brolio, F. A. Some new directions on the study of the Italian majolica pottery “a la Turchesca” known as “Candiana”. *TheMA: Open Access Research Journal for Theatre, Music, Arts*, 37–50 (2013).
18. Brolio, F. A. in (eds Mondini, S. & Guidetti, M.) 163–170 (Edizioni Ca' Foscari, 2016).
19. Bilgi, H. *Kütahya tiles and ceramics* (Suna and İnan Kira Foundation, 2006).
20. Kürkman, G. *The magic of fire and clay: a History of Kütahya pottery and potters* (Suna and Inan Kiraç Foundation / Pera Museum, 2006).
21. Crowe, Y. Kütahya ceramics and international Armenian trade networks. *V&A Online Journal* (2011).
22. Greenwood, M. *The designs of William De Morgan* (Richard Dennis, 1989).
23. Korre-Zografou, K. *Ta keramika tou hellenkou chora* (Melissa, 1995).
24. Martens, H. & Steenaert, B. *Arnhemsche fayencefabriek : 1907-1934* (Museum Voor Moderne Kunst Arnhem & Nederlandse Vereniging van Vrienden van Ceramiek en Glas, 1997).
25. Labrusse, R. *Islamophilie: L'Europe moderne et les arts de l'Islam* (Somogy editions d'art, 2011).

26. Gibson, M. in *Art, trade, and culture in the Middle East and beyond: from the Fatimids to the Mughals* (eds Ohta, A., Rogers, M. & Haddon, R. W.) chap. Colouring the Surface: a Taste for ‘Persian’ Tiles in English Domestic Architecture, 1870–1914 (Gingko Press, 2016).
27. Moraitou, M. *Iznik. A fascination with ceramics* (Benaki Museum, 2017).
28. Ioannidis, Y. *ICARO: The pottery factory of Rhodos 1928–1988* (Benaki Museum, 2017).
29. Willis, M. D. Tiles from the mosque of Rüstem Paşa in Istanbul. *Artibus Asiae* **48**, 278–284 (1987).
30. Kaiming, H., Gkioxari, G., Dollár, P. & Girshick, R. Mask R-CNN. *Proceedings of the IEEE international conference on computer vision* (2017).
31. Johnson, J. W. Adapting Mask-RCNN for automatic nucleus segmentation. arXiv: [1805.00500](https://arxiv.org/abs/1805.00500) (2018).
32. Vuola, A. O., Akram, S. U. & Kannala, J. Mask-RCNN and U-Net ensembled for nuclei segmentation. *16th International Symposium on Biomedical Imaging (ISBI 2019)*, 208–212 (2019).
33. Yu, Y., Zhang, K., Yang, L. & Zhang, D. Fruit detection for strawberry harvesting robot in non-structural environment based on Mask-RCNN. *Computers and Electronics in Agriculture* **163**, 104846 (2019).
34. Si-Mohamed, S. Automatic knee meniscus tear detection and orientation classification with Mask-RCNN. *Diagnostic and Interventional Imaging* **100** (Mar. 2019).
35. Zuo, L., He, P., Zhang, C. & Zhang, Z. A robust approach to reading recognition of pointer meters based on improved Mask-RCNN. *Neurocomputing* **388**, 90–101 (2020).
36. Toda, Y. *et al.* Training instance segmentation neural network with synthetic datasets for crop seed phenotyping. *Communications Biology* **3**, 173 (2020).
37. Schaap, E. B. *Dutch floral tiles in the Golden Age and their botanical prints* (Becht, 1994).
38. Benn, J., Marsland, S., McLachlan, R. I., Modin, K. & Verdier, O. Currents and finite elements as tools for shape space. *Journal of Mathematical Imaging and Vision* **61**, 1197–1220 (2019).
39. Kim, D., Son, S.-W. & Jeong, H. Large-scale quantitative analysis of painting arts. *Scientific Reports* **4**, 7370 (2014).
40. Elgammal, A., Mazzone, M., Liu, B., Kim, D. & Elhoseiny, M. *The shape of art history in the eyes of the machine* 2018. eprint: [arXiv:1801.07729](https://arxiv.org/abs/1801.07729).
41. Kim, M. & Kim, J. Comparative analysis of Dutch art and impressionism through an interdisciplinary method. *10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI 2017)*, 1–6 (2017).
42. Kim, M. & Kim, J. Complementary quantitative approach to unsolved issues in art history: similarity of visual features in the paintings of Vermeer and his probable mentors. *Leonardo* **52** (2017).
43. Sigaki, H. Y. D., Perc, M. & Ribeiro, H. V. History of art paintings through the lens of entropy and complexity. *Proceedings of the National Academy of Sciences* **115**, E8585–E8594 (2018).
44. Lee, B. *et al.* Dissecting landscape art history with information theory. *Proceedings of the National Academy of Sciences* **117**, 26580–26590 (2020).
45. Perc, M. Beauty in artistic expressions through the eyes of networks and physics. *Journal of The Royal Society Interface* **17**, 20190686 (2020).
46. Bell, P. & Ommer, B. in *Connoisseurship nel XXI secolo. approcci, limiti, prospettive* (eds Aggujaro, A. & Albl, S.) (Artemide, Rome, 2016).

47. Gibson, M. in *Glazed wares as cultural agents in the Byzantine, Seljuk and Ottoman lands* (eds Yenisehirlioglu, F., Böhlendorf-Arslan, B. & Kontogiannis, N.) chap. The China syndrome: a study of the origin of hatāyī in early Ottoman ceramics (Koç University Press, 2021).
48. Gray, R. D., Drummond, A. J. & Greenhill, S. J. Language Phylogenies Reveal Expansion Pulses and Pauses in Pacific Settlement. *Science* **323**, 479–483 (2009).
49. Bouckaert, R. *et al.* Mapping the Origins and Expansion of the Indo-European Language Family. *Science* **337**, 957–960 (2012).
50. Tehrani, J. J. & Collard, M. On the relationship between interindividual cultural transmission and population-level cultural diversity: a case study of weaving in Iranian tribal populations. *Evolution and Human Behavior* **30**, 286–300.e2 (2009).
51. Tehrani, J. J., Collard, M. & Shennan, S. J. The cophylogeny of populations and cultures: reconstructing the evolution of Iranian tribal craft traditions using trees and jungles. *Philosophical Transactions of the Royal Society B: Biological Sciences* **365**, 3865–3874 (2010).
52. Matthews, L. J., Tehrani, J. J., Jordan, F. M., Collard, M. & Nunn, C. L. Testing for divergent transmission histories among cultural characters: a study using Bayesian phylogenetic methods and Iranian tribal textile data. *PLoS One*, e14810–e14810 (2011).
53. Venables, B. & Ripley, B. *Modern applied statistics with S* (2002).
54. Bonhomme, V., Picq, S., Gaucherel, C. & Claude, J. Momocs: Outline Analysis Using R. *Journal of Statistical Software* **56**, 1–24 (2014).
55. Krijthe, J. H. *Rtsne: T-Distributed Stochastic Neighbor Embedding using Barnes-Hut Implementation* R package version 0.15 (2015).
56. Scrucca, L., Fop, M., Murphy, T. & Raftery, A. mclust 5: Clustering, classification and density estimation using Gaussian finite mixture models. *The R Journal* **8**, 205–233 (2016).

Supplementary Information

S1 Tables

figure	our id	origin	location	accession number
1A	HAM01	Iznik / Istanbul (1520–1540)	Harvard Museum, Cambridge, MA	1960.19
1B	BM40	Iznik (1540–1560)	BM, London	G.42
1C	VA05	Iznik (1550–1555)	V&A, London	C.1997–1910
1D	BM69	Iznik (1540–1560)	BM, London	G.67
1E	BM68	Iznik (1570–1600)	BM, London	1878.1230.497
1F	BEN11	Iznik (1570–1575)	Benaki Museum, Athens	inv. no 79.
1G	VA69	Iznik (c. 1580)	V&A, London	1645–1892
1H	BEN06	Iznik (c. 1600–1630)	Benaki, Athens	GE75
1I	—	Damascus (1550–1600)	BM, London,	1960.0519.1
1J	iVA15	Florence, Cantagalli (c. 1892)	V&A, London	304–1892
1K	iBM14	Paris, Theodore Deck (c. 1870)	BM, London,	2019.8010.12
1K	iVA70	Derbyshire, Pilkington (1862)	V&A, London,	C.203–1976

Table S1: Sources for Figure 1. BM: British Museum; V&A: Victoria & Albert Museum.

Parameter	value
number of classes	3
backbone	ResNet-50, FPN
images per batch	4
base learning rate	0.02
batch size per image	512
maximum iterations	25000
augmentation	ResizeShortestEdge, RandomFlip

Table S2: Configuration of the image segmentation task

years	blues	turquoises	greens	purples	reds	blacks
1480–1500	black–cobalt (1–2 tones)	—	—	—	—	—
1501–1520	bright–cobalt (3–4 tones)	—	—	—	—	—
1521–1540	bright–cobalt	lilac, dark	clear	—	—	—
1541–1560	bright–cobalt, pale	clear, greenish	sage, pale	dark, medium, pale, lilac	orange	outlines
1561–1620	bright–cobalt, lilac, dark	clear, greenish	emerald, dark, turquoise–green	—	bright, medium, dark–brick	outlines
1621–1700	bright–cobalt, dark	greenish	—	—	dark–brick	outlines

Table S3: The evolution of the canonical Iznik palette: 1480–1700. Each term refers to a distinct colour unless specified as multiple tones.

S2 Defining motif families

It is easy to identify ideal representatives of particular Iznik motif families, however, a comprehensive survey shows many deviations from those ideals. Iznik artesans painted many different kinds of flowers and leaves, but since they didn't care very much about botanical accuracy, it can be hard to identify which species they had in mind as they picked up their brushes. When constructing our training set we therefore devised operational definitions to help us identify motif families.

tulips

We define a “tulip” as a representation of a petaloid monocot (tulips, daffodils, lilies *etc*) flower with the following properties:

1. lateral (rather than corollar) view.
2. distal-proximal (base-tip) axis typically greater than left-right axis.
3. simple, oblate, petals.
4. sepals that do not embrace the flower or exceed its distal margin. This excludes “rose buds”.
5. a corolla with a maximum diameter midway along the proximal-distal axis rather than distally (i.e., tulip- rather than cup-shaped). This excludes other monocot flowers.
6. single flower per stem. This excludes “campanulas” which are represented as having many flowers coming from a single stem.
7. leaves, if present, are caudine and strap shaped. They are not segmented.

carnations

We define a “carnation” as a representation of a flower with many, but not necessarily all, of the following features:

1. lateral (rather than corollar) view. Carnations are frequently depicted from a corollar view, but we exclude these.
2. multiple petals.
3. lobate petals.
4. red colour.
5. calyx with a maximum diameter midway along the proximal-distal axis rather than distally (i.e., tulip- rather than cup-shaped)
6. sepals that do not embrace the flower or exceed its distal margin. This excludes “rose buds”.
7. has a pair of anthers protruding beyond the corolla.

saz-like leaves

The descriptor “saz-like” draws its name from the “saz leaf” a motif said to have been introduced into the Iznik lexicon by Şahkulu around 1535. Although the classic saz leaf is well known, we found a continuum of shapes and sizes from rather small and simple serrated leaves to the large and complex classic form. Given this we found it impossible to decide how large, how complex, and how sinusoidal a leaf must be in order to qualify as a true saz leaf. For this reason we included all leaves that meet the following features, not all of which need be present, in the class of “saz-like” leaves:

1. acicular (needle) shape, that is, with a distal-proximal (base-tip) axis greater than left-right axis
2. a rachis (“spine”).
3. serrate or lobulate margins.
4. isolated, though they may overlap or inter-penetrate or be arranged in a basal rosette (as in a dandelion).
5. twisted into a simple curve or s-shape.
6. depicted from a partial lateral, rather than a dorsal or ventral, view so that the blade of the leaf is on one side of the rachis.
7. multi-coloured.
8. If compound (composed of many leaflets emerging from a rachis), then the entire unit, rather than the individual leaflets, is considered as a “saz-like leaf”.
9. If two more leaves are fused (without a stem between them) then the entire unit is considered a single “saz-like leaf”; if separated by a stem then they are considered as separate “saz-like leaves”.
10. The following leaves are not counted as “saz-like leaves”:
 - (a) leaves attached to a stem (as in a rose-bud),
 - (b) an appendage to a larger ornamental motif such as a rosette,
 - (c) tulip leaves.

S3 Palette discovery by unsupervised learning

As an alternative to the expert-supervised palette used in our analysis we also constructed palette by means of unsupervised clustering methods alone. To establish how large the unsupervised palette should be we first sampled 1% of all pixels in all segmented objects 100 times and carried out k -means clustering on the CIELAB values of each of these samples, with $2 \geq k \leq 40$. Silhouette scores and within-cluster sum of square, WSS, sums-of-squares both indicated an optimal partition of just two colours but also a secondary optimum at $k = 8$ or 10 (Figures S1 A, B). Since this secondary optimum is close to the size of the expert palette we sought a palette of 10 colours. We then discovered the palette itself by a sequential clustering method designed to minimize the colour variation among pixels within objects and objects within artefacts. We adopted this procedure reasoning that the pixels of a given object, and the objects on a given artifact, will tend to be similar, that is, blue tulips on a single plate are probably the same blue. In the first round we clustered the pixels within each object using $k = 20$. This ensured that rare colours within each

object were captured. At this point we have a palette for each object. We then subjected these object palettes of each artefact to a second round, again using $k = 20$. At this point we have a palette for each artefact. This ensured that all similar objects within artefact will have the same colours. These artefact palettes were then subjected to a final round of clustering using $k = 10$ to generate the final unsupervised palette. The colour values of this final palette were, then, assigned to all pixels of all objects. Figures S1C. and D shows a comparison between the expert-supervised and unsupervised palettes. They occupy the same colour space but divides it slightly differently giving. The unsupervised palette, for example, has three blues instead of one; two reds instead of one; one green instead of two and no turquoises or purples. It appears that our unsupervised palette is missing some rare colours. Since these colours are important we use the expert-supervised palette in all analyses.

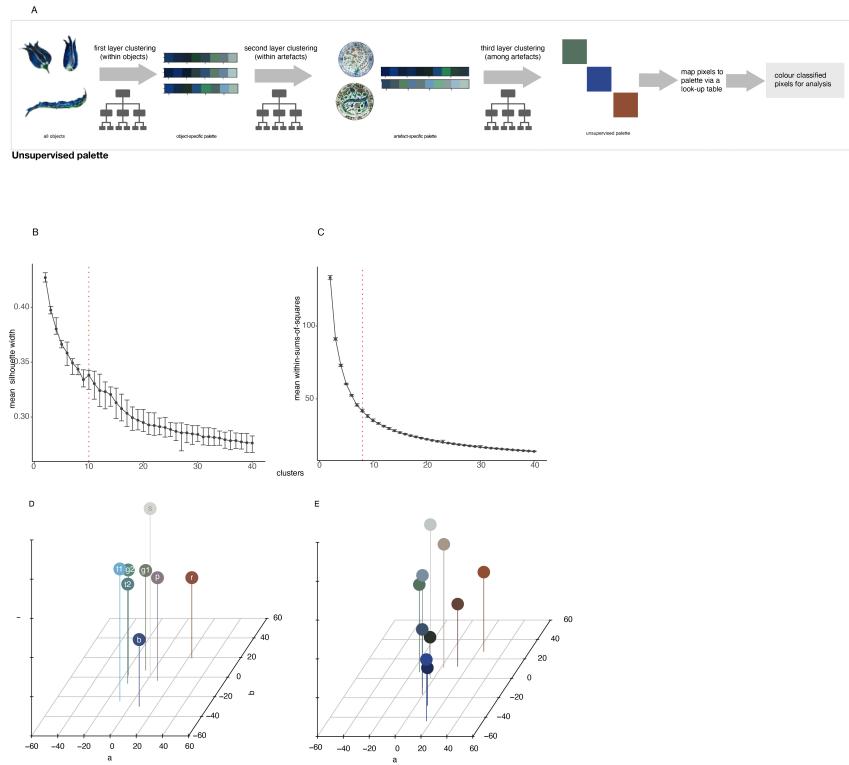


Figure S1: **A.** Flowchart showing the method of unsupervised palette discovery. **B. & C.** Estimating the best k for k -means clustering on the CIELAB values of object pixels. Means of silhouette scores and within-cluster sum of squares with 95% confidence intervals. The elbow of the within-cluster sum of square curve occurs at $k = 8$, and a secondary optimum occurs in the silhouette scores at $k = 10$. We chose the later for the size of the unsupervised palette. **D.** Expert-supervised palette as shown in Figure 4 compared to **E.**, the unsupervised palette. Note the dull colours and lack of turquoise and purple in the latter.