

AI and Mathematics Interaction for a New Learning Paradigm on Monumental Heritage



Álvaro Martínez-Sevilla and Sergio Alonso

1 Introduction

The purpose of this chapter is to show how AI can help, from fields such as computer vision or deep learning, to understand and explore mathematical concepts. We present two AI-based systems that incorporate interaction with reality in an environment that eases the use of mathematical concepts and applications in a highly motivational field for its cultural and historical load: the monumental heritage. To do so, in the rest of this section we present how digital technologies have been introduced into mathematics education and how the intrinsic mathematics in monuments conform an interesting tool to explain and reinforce mathematical concepts. In Sect. 2, we present the main forms of mathematical reasoning and how AI offers a new way of reasoning worthy of investigation. In Sect. 3, we present MonuMAI, an AI-driven environment designed to help the general public to do monumental analysis. Not only do we present the tool but also we describe some of the citizen science methodology that has been developed with it and the problems that arose during the development and use of the system. In Sect. 4, we present a novel approach to construct an automatic geometrical model of architectural façades. To do so, we have to overcome some issues as the rectification of images and photographs in order to get an image without perspective distortions that can be later automatically analyzed to discover geometrical properties as symmetries. In Sect. 5, we briefly describe some of the educational experiences

Á. Martínez-Sevilla (✉)

Research Institute on Data Science and Computational Intelligence, Universidad de Granada, Granada, Spain

e-mail: asevilla@ugr.es

S. Alonso

Software Engineering Department, Research Institute on Data Science and Computational Intelligence, Universidad de Granada, Granada, Spain

e-mail: zerjioi@ugr.es

that we have carried out using the presented tools. Finally, in Sect. 6 we provide some indications on the future works that we expect to develop.

1.1 Mathematics Education: Reality, Models, Computation and Solutions

Mathematics education has been a unique and exceptionally powerful way of teaching our young people. It usually takes problems from reality and models them by determining objects and relationships between them. Finally, it develops a rigorous method to combine these objects and relationships in order to obtain answers to the problems posed, making the solutions explicit in the terms in which we had initially taken them from our real context. To address a set of problems expressible under the same model, it develops theory and methods, in which finally mathematics education will have to be able to transmit the model and the general solution methodology. Of course, it is also important to have the ability to recognize reality and fit it into one of the given models (or a variant). Inspiration from reality is a central source of motivation. The *example* is the most evident proof that our mathematical and model development works on concrete cases and deepens the appreciation of reality in a systematic and clarifying way.

Therefore, the connection with reality is an inalienable way of teaching mathematics. This occurs more intensely at the first levels, in primary and secondary education. As mathematics education advances at higher levels, this connection becomes less necessary. At university levels, it provides students with methods and developments, which they have to be able to put into practice in contexts outside mathematics itself.

Perhaps the most time-consuming part of mathematics education is the one which focuses on managing the model to obtain the solutions. That is, developing correct computational mechanics that proceeds without inappropriately mixing objects or relationships between them. Fortunately, the algorithmic nature of the mathematical reasoning process has led to automating many of these processes with the help of digital technology in recent decades. This has allowed us to free ourselves from having to calculate on each occasion in very repetitive processes which are not really interesting or enriching for the understanding of the world that we intend for our students.

This allows us to focus on the concepts, on reality and its conversion into a model, on the identification of its elements and on the correct interpretation of the obtained solution. Mathematics education must be able to devote adequate time to each part of the mathematical learning process, giving perhaps greater importance to concepts and reality, and perhaps less to calculation and the use of digital support tools. In fact, one of the objectives of technology has been to simplify these processes and to make them increasingly intuitive and less dependent on specific programming skills, except in very specific and specialized cases of problem-solving at a high educational level or in specialized training.

1.2 *Computational Support in Mathematics Education*

The first portable electronic calculators, with low power consumption chips, appeared in 1970, and they represented a revolution, a great leap from the mechanical machines of arithmetic calculation that were highly developed toward the middle of the last century. With only a disposable battery and an integrated circuit, calculators are capable of simple and direct computations. The low cost and portability made them become a usual element of the school portfolio in the following years (particularly since 1976). They have been incorporating more and more calculation functions, such as trigonometric or statistics. However, some discussion persisted among different educational systems on the desirability of keeping in teaching the ability to perform the accounts with pencil and paper or with a calculator.

This first generation school machines reach what we could call the second generation with the incorporation of symbolic abilities. In 1987, HP launched a model with symbolic equation solving capabilities. In the following years, improvements and the incorporation of more symbolic capabilities followed, running in parallel with improvements in hardware, where PDAs and Pocket PCs were added to the repertoire of calculation devices. Also, software has improved in order to process text and spreadsheets. They are also connectable devices, via cables, infrared beams and other technologies, which makes them more versatile.

Parallel to these advances, countless symbolic calculus packages for personal computers emerged. However, the rigidity and complexity of their programming made them a tool for professionals or high educational levels, and they were generally rejected by mathematics teachers in primary or secondary schools.

We could say that the third generation starts with the presentation of GeoGebra (2021) by Markus Hohenwater in 2002. It is a software program that aims to combine dynamic geometry with symbolic computation, allowing intuitive and easy handling of the provided functions. GeoGebra is widely accepted within the educational community and has propitiated the creation of a large-scale network of collaborators who create applications with a didactic approach. In recent years, GeoGebra has improved its functionalities and even created specialized versions for the web and smartphone devices. Thus, it is becoming a very used tool for mathematics education due to its ease of use, flexibility and pedagogical adequacy, maintaining the possibility of understanding mathematical concepts, while facilitating their calculation.

The devices that we have called first and second generation have a common characteristic: they make calculations, either numerical or symbolic, but they do not interact with reality. This interaction appears with GeoGebra, with the ability to load any images and construct dynamic geometry on them. This leads to an innumerable source of didactic applications on concrete problems, taken from immediate everyday life. GeoGebra intends to expand these capabilities with the introduction in 2018 of the GeoGebra AR version, which runs on a smartphone and uses the device's camera video stream to build mathematical objects over it. Although still limited in the ability to load arbitrary Geogebra-like files, it is a big step in the integration of the

computing device—here mainly to plot mathematical objects—with the interaction with the world.

However, we believe that the potential of GeoGebra in its interaction with reality is largely unexplored. In Botana et al. (2020), we propose the creation of an extension for GeoGebra that is able to interact in a deeper way with reality, through the use of a wide range of sensors of a smartphone: camera, accelerometer, compass, and positioning system (currently GPS), in such a way that allows measuring distances, angles, speeds, geographical positions, directions and so on. GeoGebra would thus take advantage of the large number of sensors in a smartphone to integrate the reality-model interface in the calculation device, something that started motivating mathematics education and which is still difficult to achieve by computer systems.

The fourth generation of these digital devices to facilitate mathematics education are Artificial Intelligence (AI) systems. They can use the computational capacity and connectivity of smartphones while integrating them with tasks such as cloud computing, Internet of Things, and access to databases, and thus constitute a whole universe of possibilities for interactive calculation against reality. The coming years will surely bring us an exciting path in the development of this integration to perform increasingly complex tasks in a simpler and more intuitive way. The use of expert knowledge to approach and integrate other scientific or humanistic subjects, analyzing solutions and proposing application scenarios or intensifying the use of technology are some of the desirable features in what we call STEAM (science, technology, engineering, arts and math) education. These features will undoubtedly come hand in hand with the creation of AI-based systems. Particularly, in the following sections we will describe two of these systems that incorporate interaction with reality.

1.3 The Mathematics in Monuments: A Unique Form of Mathematics Education

One way to bring students closer to mathematics is by connecting it to everyday life. Not only does it allow a surprising and close application scenario but it is also very accessible for special sessions and projects in the classroom and outside. In fact, some technological applications have been developed in this sense specifically dedicated to supporting the educational actions with mathematics in the scenario of a city.

From this point of view, mathematics linked to the monuments have a special place: to the rich mathematics contained in monuments, we must add their interdisciplinary possibilities and the motivation they generate. Their mathematical properties can be presented together with their historical-artistic aspects as well as other scientific aspects linked to monuments. They offer great opportunities to study through technology, making this field one of the most promising in mathematics of STEAM projects and orientations.

In this sense, we have already developed a methodology (Martínez-Sevilla, 2017a) and mathematical content resources (Martínez-Sevilla, 2017b, 2020) as well as technological ones (Botana et al., 2020) that allow an approach to the field of teaching mathematics based on monuments. Beyond simple contents, mathematics in monuments offers the rich complexity of the intricacies of the relationships in which artists, architects and geometers have come together to design the appearance of buildings to last for the future and to transmit a symbol to future generations. These symbolic aspects are perhaps the most attractive and motivating in the use of monuments in mathematics education. It is so that on occasions, they have even made it possible to provide a new sense of historical-artistic interpretation mediated by new research based on mathematics (Martínez-Sevilla & Cruz Cabrera, 2021). Moreover, the use of mathematics just for functional or decorative aspects, where calculus and geometry usually play an important role, has allowed numerous teaching resources based on them.

Naturally, this scenario could not be unaware of the great development of Artificial Intelligence in the last decade (Fiorucci et al., 2020). Both as a tool for monumental analysis and as a proposal for an interdisciplinary educational approach which manages resources, technologies and modes of reasoning typical of a recent discipline, but with increasing depth in current teaching. Within this framework, we have developed the MonuMAI tool (Lamas et al., 2021) as an integrated and growing system to approach the monumental environment through AI and mathematics.

2 Forms of Mathematical Reasoning: Deduction, Induction, and Abduction

In this section, we introduce the classical forms of mathematical reasoning (deduction and induction) and present how AI offers a different (but interesting) way of reasoning (abduction).

2.1 Classical Forms of Reasoning

Among the most defined forms of mathematical reasoning, there are two that are usually applied and taught in mathematics: **deduction** and **induction**.

In **deductive reasoning**, a conclusion is obtained from a finite number of previously established hypotheses or premises. A series of rules, called inference rules, intervene in its definition, which we will use in our reasoning. Thus, we would have, as a more formal definition, that *a deduction is a finite set of steps, in which in each one we use either one of the hypotheses (or axioms added to these) or a formula obtained by applying one of the inference rules to two of the previous steps*. In this way, the first two steps must necessarily be hypotheses or axioms. The last step is

the so-called conclusion, deduced from the set of hypotheses. From this mode of reasoning, science has elaborated the *hypothetical-deductive reasoning*, where the observation of reality and the experimentation play their role to verify the system of hypotheses and conclusions.

The other mode of reasoning, that is, **inductive reasoning**, works in a very different fashion. In general, in inductive reasoning we go from the particular to the general. Induction studies the properties that a certain object or sequence fulfills in order to extend or generalize this property to a wider family in which essentially the same norms are fulfilled. It is a usual way of learning in children and even in the daily reasoning of our life. As an example, we use this kind of reasoning if we observe on multiple occasions that after dark clouds rain comes. We will then tend to deduce that rain comes from the presence of dark clouds when this may not really be the case: the consequence of dark clouds may also be snow, or no rain due to other meteorological variables.

However, in classical mathematics, violations of a general formulated rule are not tolerable. That is why they use the most far-reaching format of induction reasoning, the so-called *Complete Induction*, which in its strongest form can be formulated as follows.

If $A \subset N$ and

1. $0 \in A$
2. for any $n > 0$, $\{0, 1, \dots, n - 1\} \subseteq A$, then $n \in A$

then $A = N$.

This is the so-called **Second Principle of Induction**. There is another, with a weaker formulation, known as the **First Principle of Induction**. Both principles are equivalent and constitute an axiom of natural numbers. In fact, they are valid for any set that is a *Well Order*, that is, in which each subset has a minimum element. Moreover, the demonstration of this property can be deduced from the Principle of Induction.

Therefore, this is a form of mathematical reasoning that allows to extend a property to the whole set of natural numbers as long as that property has the first element that fulfills it and that given an arbitrary number that fulfills it, we can verify that the following one also fulfills it (weak formulation).

Inductions can also be applied to two or more variables. All of them are ways to generalize a demonstration in an automatic way to a larger set than the starting one. They form the basis of the logical work with some interesting mathematical objects as recurrences or successions.

2.2 *AI and Its Forms of Reasoning*

However, it is not easy to fulfill the deduction and induction requirements in practical applications. Classical mathematics, which is the one that is mainly taught in primary and secondary schools, has established a type of universal truth based on binary logic,

whether it is fulfilled or not, and with no other options to consider. Usually, however, we are presented with situations that are not so clear-cut.

For example, reasonings in which we can use the word *probable*. Estimates about the occurrence of some event, which may trigger a deduction, but only in a probabilistic way. Even the application of some rules may be affected by such a measure of probability. This type of reasoning is frequently found in daily life and in scientific applications, but it is not usually trained in mathematics lectures (Batanero et al., 2005, 2016). On many occasions, we talk about probability, and we estimate it with measurements, but usually disconnected from the world of logic, without incorporating it within a context of inference rules. Artificial Intelligence basically works with this kind of reasoning and that is why familiarity with it should be the first step to acquire this argumentative ability.

Therefore, a new mathematics education is necessary, not only to provide validity to the statements obtained by deduction or induction, but also to those obtained by **abduction**: the method of reasoning in which probability is incorporated in the premises or the rules of inference and also in the conclusion.

There are several types of logic that incorporate the management of uncertainty in different forms and for different purposes: *modal logic*, *probabilistic logic* or *fuzzy logic*. In them, not only probabilities are quantified for a premise or inference rule but also imprecise concepts are handled to indicate different types of quantities in everyday life. For example, in modal logic we can use statements such as “it is necessary that” or “it is possible that”, in probabilistic logic, we can use “probably (with a crisp measure)” and in fuzzy logic, we can use linguistic terms (qualitative quantifiers) such as “many” or “surely”.

Mathematics education should make an effort to incorporate these types of logics that are capable of managing the uncertainty of an answer in a generalized way. Although modal logic can be understood as a formal and didactic step in that management and fuzzy and probabilistic logics as technical tools that allow reasoning in environments with partial information and uncertainty, it is important to incorporate them with the appropriate depth to each educational stage as well as to provide examples and applications that practice with them. Today’s world is already working through AI, and we can find some examples of the interest that it brings forward to mathematics education (Chassignol et al., 2018; Gadanidis, 2017). It is used as a base and therefore an adequate understanding of its possibilities and results will only be possible with an education that incorporates adequate conceptual models of reasoning.

One of the AI-based technologies that can be successfully used to illustrate the power of AI and its way of reasoning is pattern recognition, a set of techniques aimed to recognize in an automated way patterns and other regularities in data. Please note that pattern recognition differs from pattern matching in the sense that the latter is aimed to find *exact* coincidences in the data while the former pretends to find softer regularities in the data (not exact coincidences but with a certain degree of flexibility). As an analogy, we can say that pattern matching is as strict as inductive and deductive reasoning processes, while pattern recognition follows a more abduction-like form of reasoning.

Particularly, visual pattern recognition takes as input images and tries to find characteristics in those images to classify them (or to classify sub-elements in the images). Due to its visual nature, it makes for an appropriate tool to use with scholars.

3 MonuMAI: An AI-Driven Environment for Monumental Analysis

As we have mentioned before, monuments can concentrate a large part of the mathematical knowledge of their time, as well as valuable historical and artistic information. This is particularly intense on monumental façades, so we will focus mainly on them. The analysis of a façade requires a lot of expert knowledge:

- different styles usually share some similar architectural elements,
- monuments usually have a mixture of styles and periods, and
- some are not executed under canonical criteria, but with singularities.

All these factors, with the addition of the deterioration and alterations of the monuments as time passes, make this kind of analysis a quite difficult task.

In order to facilitate it and, at the same time, to create a teaching support tool that can serve in an interdisciplinary way in subjects of mathematics, art and history, we have created MonuMAI (Monuments with Mathematics and Artificial Intelligence). MonuMAI is an interdisciplinary project of research, education and scientific dissemination, joining researchers from Math, AI, general Computer Science, and History of Art, altogether scientific communicators and educational advisors. It is a STEAM project which has as its main objective to deal with science and art by means of technology.

MonuMAI, born in September 2018, is a project of the Research Institute on Data Science and Computational Intelligence¹ of the University of Granada² and the Descubre Foundation.³ You can check out more information and details about it at <http://monumai.ugr.es/> (Fig. 1). The related app can be downloaded from both the AppStore (iOS) and Play Store (Android), or directly from the links on the project page.

3.1 *MonuMAI Dissection: How to Classify Monuments Using Deep Learning*

MonuMAI consists of three main blocks:

¹ <https://dasci.es/>.

² <https://www.ugr.es/>.

³ <https://www.fundaciondescubre.es/>.



Fig. 1 The front page of the MomuMAI web page and the splash screen of its app

1. A monuments dataset;
2. A deep learning pipeline;
3. A mobile app to integrate the previous ones.

In the following we describe the three parts:

Monuments dataset: An extensive dataset (selected and annotated by experts in Art History) has allowed us to train the algorithm for the recognition of up to 15 key architectural elements, easily recognizable by their geometry (Fig. 2). These in turn will allow us to distinguish up to 4 from the most important artistic styles in Europe: Renaissance, Baroque, Gothic and Hispano-Muslim. This dataset is a database of labeled images of monumental façades. Such a database did not exist to date. The MonuMAI dataset includes 6650 tagged images of the 15 key elements. The dataset is the basic component, a result of expert knowledge, on which MonuMAI bases its learning through its pipeline.

Deep learning pipeline: The MonuMAI deep learning pipeline's main purpose is the detection of key elements in an image using the MonuMAI Key Element Detection (MonuMAI-KED) model and also the classification of artistic styles.

MonuMAI-KED is based on a novel taxonomy of monumental heritage (MonuNet), specifically created for this task (Fig. 3). This taxonomy incorporates a classification of styles according to the recognition of the previous 15 key architectural elements. The classes cover arch, structural support objects and decoration, horseshoe arch, lobed arch, flat arch, pointed arch, ogee arch, trefoil arch, serliana, triangular pediment (or pointed pediment), segmental pediment, gothic pinnacle, rounded arch, lintelled doorway, porthole, solomonic column and broken pediment (Fig. 2).

So, how does MonuMAI classify the style of the monuments? The 15 architectural elements allow us to establish some identification criteria for styles. For example, the Gothic style is characterized by the use of ogee arches, pointed arches, gothic pinnacles and trilobed arches. On the other hand, the Hispano-Muslim style is associated with horseshoe or lobed arches while the Baroque style uses broken pediments or solomonic columns. In this way, a pseudo tree appears, to classify styles: the 4 child nodes of the root are determined based on the descendant nodes that are reached. We

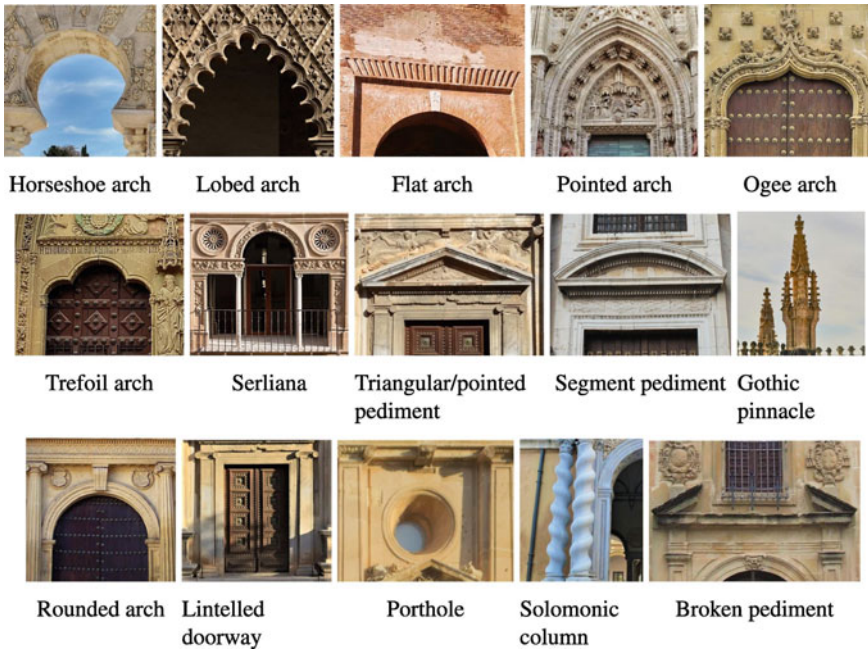


Fig. 2 The 15 architectural elements for MonuMAI-KED

would like to remark that *MonuNet* is not really a tree, but actually a graph with a few cycles. This is so because some styles share elements (descendants nodes), such as the rounded arch or the porthole for the Renaissance and Baroque styles.

This is also an interesting matter, since in education and some simplified scenarios binary or n -ary taxonomies are used but usually in the form of a tree, where there are no cross characteristics between several predecessor nodes. However, reality is not so perfectly separable, and usually establishing a realistic taxonomy implies having to classify the same object in several categories and having to go to secondary categories to make a proper classification. In fact, we can find this issue in numerous everyday examples and it appears in the field of science as well. For example, although Darwin’s well-known “tree of life” drawing is really a tree-like graph, we know that the complexity of life in many species requires a continuous modification of its taxonomy tree labels and branches. This is done to rearrange the new knowledge, and sometimes this connects leaves (or nodes) in some biological species (Dale, 2017).

Thus, it is interesting to note that *MonuMAI*’s classification may sometimes output a double-class as Renaissance/Baroque. The only possibility to discern those two styles is finding in the façade other recognized key elements which are typical of only one of the styles. If this is not possible, the uncertainty in the classification will remain. This is not a bug in the application, but a singularity of the artistic classification. Moreover, the experts themselves are sometimes unable to distinguish



Fig. 3 MonuNet: Taxonomy in the form of a rooted pseudo tree for style recognition

one style from another except when using added information such as the date of construction or the author of the design. Classification in artistic-historical styles is a matter that is far from being completely finished in the state of the art of this social science.

Mobile app: The MonuMAI app uses the two previous elements to operate in real-life conditions. Our tests have shown that the MonuNet architecture and the detection model provide excellent results even in harsh real-world conditions such as side shot photographs, noise or blur in the image.

The app offers the interaction between the user and our image classification process. The user, once registered, can take an image of a monumental façade and select a region of interest on his mobile device. Once this is done, the image is sent to our server where it is processed by MonuMAI-KED which uses a Convolutional Neural Network (CNN) to analyze positive identifications (with enough probability) of the

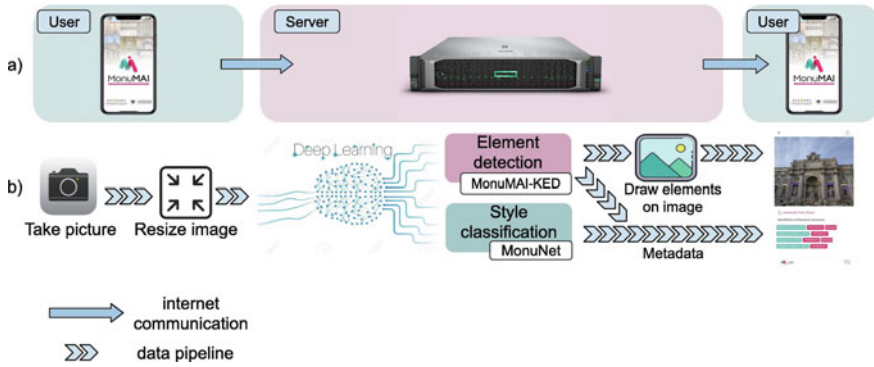


Fig. 4 MonuMAI deep learning pipeline. **a** Sequenced communications flow. **b** Image processing status in each stage



Fig. 5 Application examples in the user app: Left: Royal Court of Justice of Granada (with framed detected elements); Right: Granada Cathedral (identification and labeling of elements)

key elements detected in the image. I will also draw a frame around each of the analyzed regions for each key element on the image. Based on the detected key elements, MonuNet will offer a style classification proposal—probabilistically quantified—as metadata added to the image. Finally, the annotated image will be returned to the user. The whole process is shown in Fig. 4.

The output resulting from applying MonuMAI-KED and MonuNet to an image is then displayed in the user’s app (Fig. 5).

For the operation of MonuMAI-KED, we use a two-step strategy. First, a selective search method that is implemented as a fast region-based CNN makes a proposal

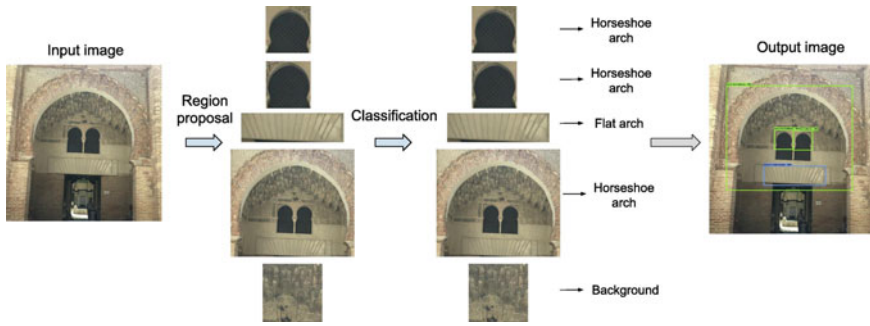


Fig. 6 MonuMAI's process for recognizing key elements in an image

of candidate regions in each input image. The search for candidate regions is done through regression learning, considering the visual features and region shape of the selected elements. In the second step, each region is classified by a CNN as an architectural element or background. For the regions classified as key elements, the architectural information is added to the monumental image. The process is shown in Fig. 6.

3.2 *The Citizen Science Methodology on MonuMAI*

One of the novel points in the methodology is that the user actively participates in the improvement of the dataset through their added knowledge. It is done applying the Citizen Science methodology: citizens participate in the project and contribute with their collaboration to achieve scientific purposes. In our setting, the user can contribute in the improvement of the scientific results offered by MonuMAI. After classifying and returning the image to the user, it is asked for their opinion regarding the classification offered, to which they can respond with “I agree”, “I don't know” or “I don't agree”. MonuMAI will value these opinions to add them to the dataset in cases where it has confirmed the offered classification, or it will lower the probability in cases of disagreement. The dataset therefore improves with its use, thus appreciably increasing the precision in the responses of the MonuMAI system.

Citizen Science has shown itself as a powerful tool for doing science in recent years. The European Union even has a Research Program based on it: Science With And For Society.⁴ Moreover, in terms of scientific possibilities, not only for participation but also for the addition of knowledge distributed in society, the case of Deep Learning and Citizen Science is a special pairing. Where scientists cannot reach, Citizen Science can reach.

⁴ <https://ec.europa.eu/programmes/horizon2020/en/h2020-section/science-and-society>.

Until now, previous existing models were based on machine learning and computer vision techniques, which require a high degree of supervision and manual intervention. These techniques produce models with little scalability, that is, the possibility to add new elements for the detection or artistic styles. Furthermore, their operation in real environments is usually difficult due to the presence of noise and spurious elements in the image. However, they usually offer satisfactory results. Our approach in which we use machine learning through deep learning techniques, supported by a Citizen Science methodology, is a better way to increase and improve the dataset, and thus recognition.

3.3 Learning and Mistakes

How does MonuMAI learn? The implemented learning model based on deep learning technologies is not the one employed in usual mathematical reasoning, where deductive or inductive processes are applied to achieve the final decision. Deep learning works as a computational brain, and its learning model is similar to the one employed by children: it reasons by analogy based on the accumulation of cases. It takes into account the context and infers knowledge—in our case probabilistically—to be able to decide among the different architectural styles after the inspection and detection of key elements. From the dataset of annotated images, it learns the differential characteristics of each architectural element. The algorithm then is able to search over every new image and process all the characteristics it has learnt and thus to recognize elements that have similar properties, even if it has never seen them. In this sense, MonuMAI's learning process is intelligent behavior.

At this point, we would like to point out that even intelligent behavior makes mistakes: For example, MonuMAI did learn that in the context of horseshoe arches, bricks are usually associated in their appearance, forming part of it, or in the decoration of its flat lintel (whenever it exists). Thus, if it found a rounded arch from the Mudejar style (post-fifteenth century style that takes some of the earlier Islamic construction techniques, very frequent in Spain), then MonuMAI used this context association and erroneously determined that the arch was a horseshoe arch and thus concluded that the architectural style was Hispanic-Muslim, when in reality it corresponded to a Christian church.

3.4 A Critical Look at MonuMAI: How to Troll the System

To test the abilities of MonuMAI and to check the kind of errors that it can produce, we propose a challenge to the students who use it (from an educational point of view): Can MonuMAI be trolled? Can the user input images that clearly are not related to architectural monuments and make its inner algorithms to fail and recognize elements and even determine a particular architectural design?

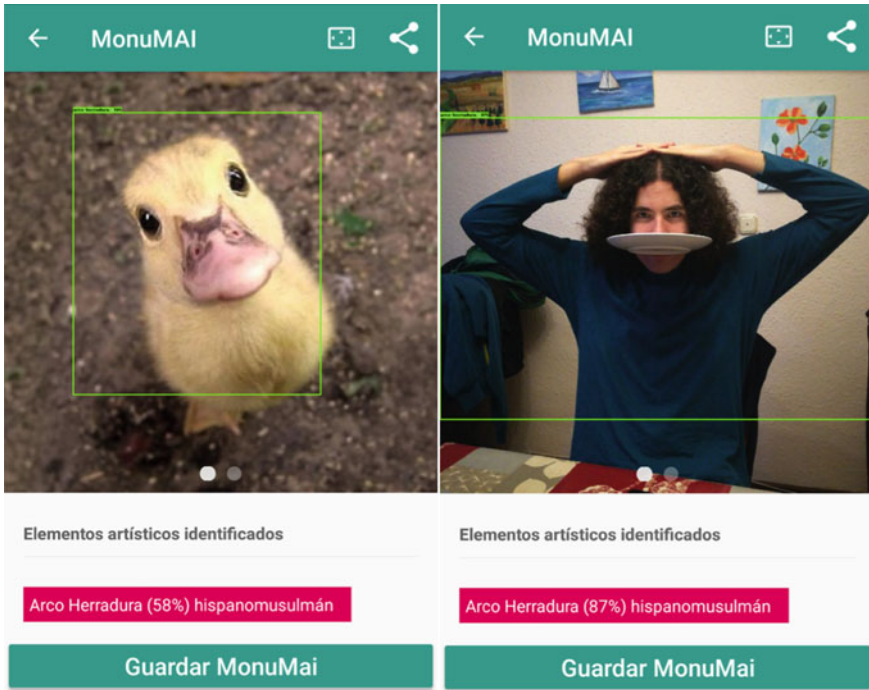


Fig. 7 Some images that deceive the MonuMAI AI

The students were pretty fast responding to the challenge. They found some example images in which MonuMAI only recognizes “shapes in their immediate context” and not the overall appearance of the image. For example, in Fig. 7 (left) the image depicts a duck. However, for MonuMAI it is not a duck but a horseshoe arch. The head with a notch on the neck, and the surrounding darker brown color (similar to horseshoe arches) makes MonuMAI infer a small probability (58%) for that architectural feature. In Fig. 7 (right), a student tried to fool the system (and succeeded): MonuMAI produced a high probability (87%) of the image depicting a horseshoe arch.

These funny examples demonstrate that the creativity of a motivated student knows no limits. Moreover, these examples offer a great opportunity to discuss with the students how a particular AI (like the MonuMAI one) may be very good in a particular task (recognize architectural styles) but it may miserably fail when faced with examples that are completely out of its scope: it is important to remark that the MonuMAI AI has never faced an image of an animal or person and therefore it cannot recognize and discriminate against them. On the other hand, the human brain has a much more general intelligence that has previously faced much more images and concepts and therefore is much more capable of differentiating against those malicious examples. However, not all humans are able to distinguish among different architectural

styles. To summarize, MonuMAI is good at classifying architectural elements and style over photographs of architectural façades, but it can say nothing correct about photographs of other different subjects.

4 Toward the Construction of an Automatic Geometric Model (AGM) of Architectural Façades

Architectural façades are a great example to understand how geometry has been extensively used in different arts and monuments. The use of particular proportions, symmetries, repeating patterns and so on can be discovered in a great variety of building façades.

Moreover, AI techniques can be used to ease the analysis of those mathematical properties and even to automatically detect the geometrical properties of the different elements in a façade. We have already covered how an AI-based mobile app can detect different art styles. However, that kind of application could be expanded not only to recognize that styles but also to automatically detect the different geometrical constructs that have been used when designing and constructing a façade.

In the following, we describe an approach to develop such a tool that analyzes and constructs geometrical models of architectural façades based in different AI techniques and computer vision technologies. This approach assumes that all the analysis must be done over a single photograph (probably taken with a mobile phone or typical camera) and without any other particular knowledge about the target. This assumption is a quite restrictive one as many advanced technologies as photogrammetry (Taddia et al., 2020) (construction of a 3D model from several pictures of the subject taken from different points of view), 3D scanning (Wojtkowska et al., 2021) and so on could provide much more precise results. However, we do not take into account these techniques and tools as they are not usually available for the general public.

4.1 *The Problem of Perspective in Photographies*

One of the first problems in order to automatically analyze a façade that has been photographed is the effects of the perspective (Soycan & Soycan, 2019). These perspective effects produce that the proportions and angles presented in the façade are not maintained in the photograph (which is a 2D representation of a 3D object). Here, it is important to remark that the human brain has evolved to be able to overcome those effects and can usually infer some geometrical properties in the photograph. Let's take as an example Fig. 8 where 2 buildings are depicted in a synthetic render. It is obvious that from the perspective chosen some of the elements in the façades are not represented at the same scale (the green circles in Fig. 9) and are deformed (the green circles are depicted as ellipses) and some parallel lines (the red ones) are not really parallel in the photograph. However, our brain can easily overcome those

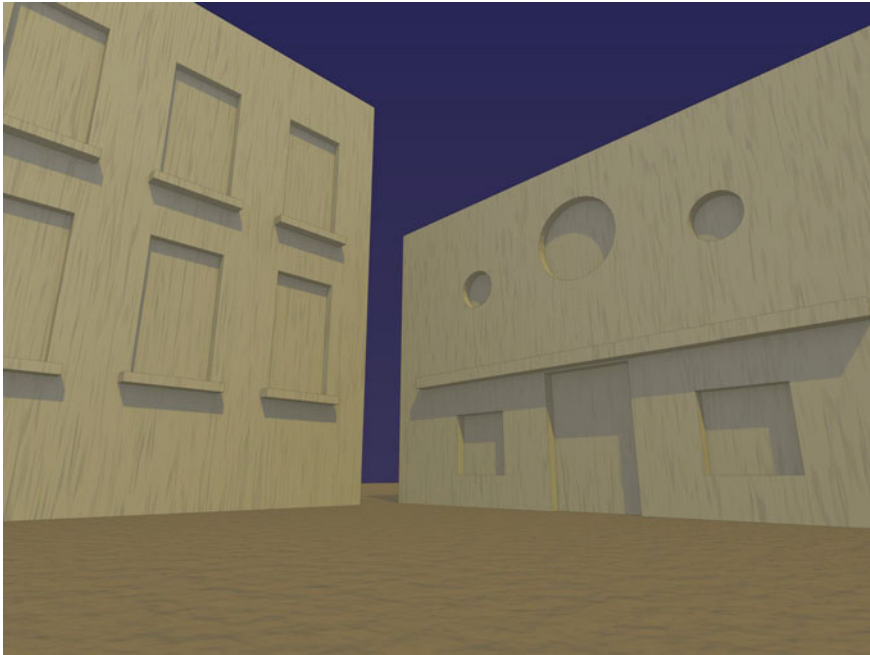


Fig. 8 Synthetic render of two buildings

difficulties and can instantly recognize that the ellipses represent circles of the same size in the façade and that the red lines are parallel ones in the real world. We think that eliciting these issues regarding perspective effects in traditional perspective drawing and photography is a very interesting topic as they can remain unnoticed by the public due to the good job that our brain does interpreting drawings and photographs.

In fact, to partially avoid those perspective effects, it would be necessary to make the photograph in a very particular scenario: we need to use specialized perspective corrected lenses, shoot at an appropriate angle in which the sensor of the camera is parallel to the façade plane and as close to the center of the façade both in the horizontal and vertical axes (which is usually not possible as the photographer cannot usually put the camera at the appropriate height). In Fig. 10, we depict the previous façade simulating those ideal conditions and we can see that the effects of the perspective are minimized.

4.2 Correcting Perspective Issues in Façades

Fortunately, today we can use computers in order to apply transformations on photographs to rectify the perspective-related issues. Image rectification consists of a

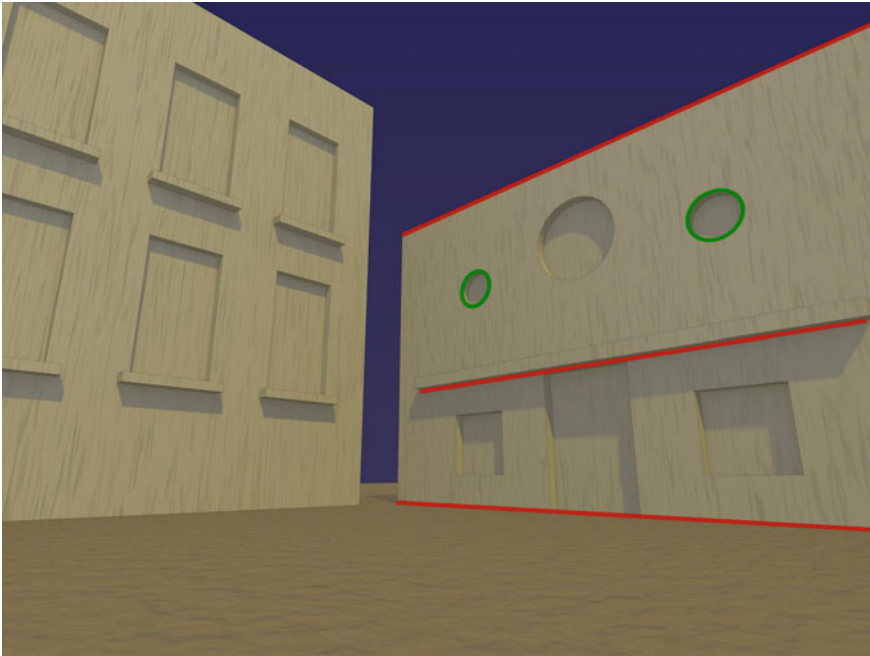


Fig. 9 Perspective distortions over elements in the façade: green circles are depicted as ellipses with different sizes, and parallel lines (in red) are depicted as not parallel

transformation process in which the original image is projected onto a different plane (in our case, a plane that makes the façade plane and the sensor of the camera to be parallel). In order to do this rectification, a projective transformation must be applied. If we know the camera angles with regard to the façade plane it is possible to compute the precise projective transformation that should be applied Soycan and Soycan (2019). These camera angles are usually called *yaw* (rotation around the vertical axis), *pitch* (rotation around the side-to-side axis) and *roll* (rotation around the front to back axis). For example, in Fig. 8 the angles of the camera were set as $yaw = 45^\circ$, $pitch = 15^\circ$ and $roll = 0^\circ$.

In our approach, we assume that we do not know those angles. However, we must note that many cameras nowadays (specially the ones integrated into mobile phones) have built-in sensors to measure that angles. Having a reasonable approximation of those rotations could be very helpful in the rectification process.

An approach to compute those angles (and even different parameters of the lens and camera) is offered in some panoramic creation software such as Panorama Tools (2007). One of the modules of this software (PTOptimizer) can take as an input a photograph and a list of lines which are identified as vertical and horizontal lines of the façade that we want to use as the rectification plane. From this information, the program executes several optimization algorithms and as a result

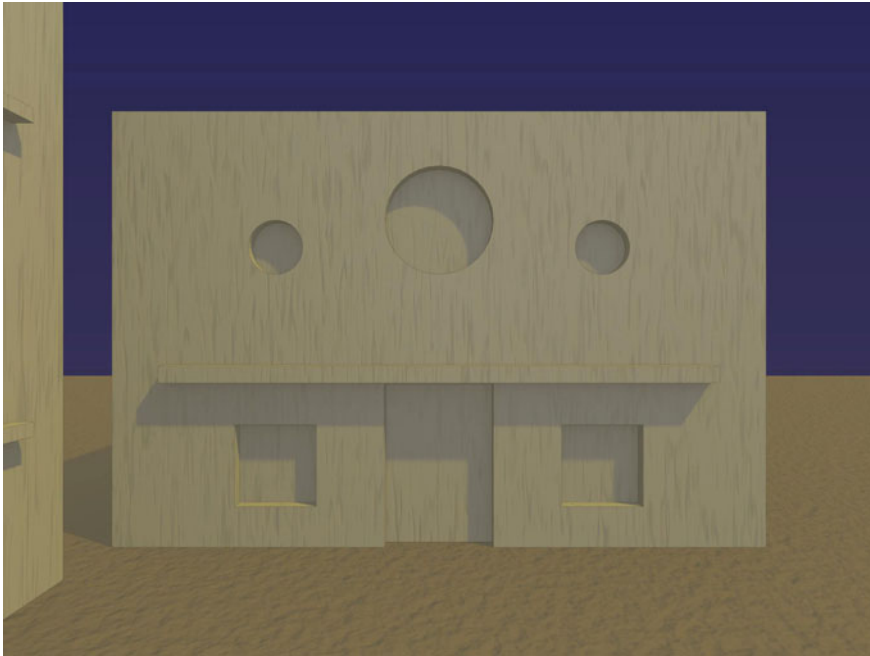


Fig. 10 Using a particular lens configuration and carefully choosing the angle and camera position, the perspective effects are minimized

provides an estimation of the yaw, pitch and roll and also the camera lenses' focal length.

4.3 The Hough Transform to Identify Basic Geometries in an Image

Therefore, in order to obtain the camera orientation we need to identify vertical and horizontal lines in the façade that we want to rectify. To do so, we propose the use of the Hough transform (Duda & Hart, 1972; Ballard, 1981). This transform allows detecting analytically defined shapes. In our case, the shapes that we want to detect are straight lines and thus, to do so, the image in which the transform is going to be applied is usually filtered by means of a Canny filter (Canny, 1986) which emphasizes all edges (abrupt changes in illumination) in the original image.

Once those edges have been emphasized, each pixel in the image “votes” in a parameter space matrix. In the case of straight lines, the parameter space is a two-dimensional one (the angle θ of the normal of the line and its algebraic distance ρ from the origin). Once the voting has finished, the algorithm just searches for local

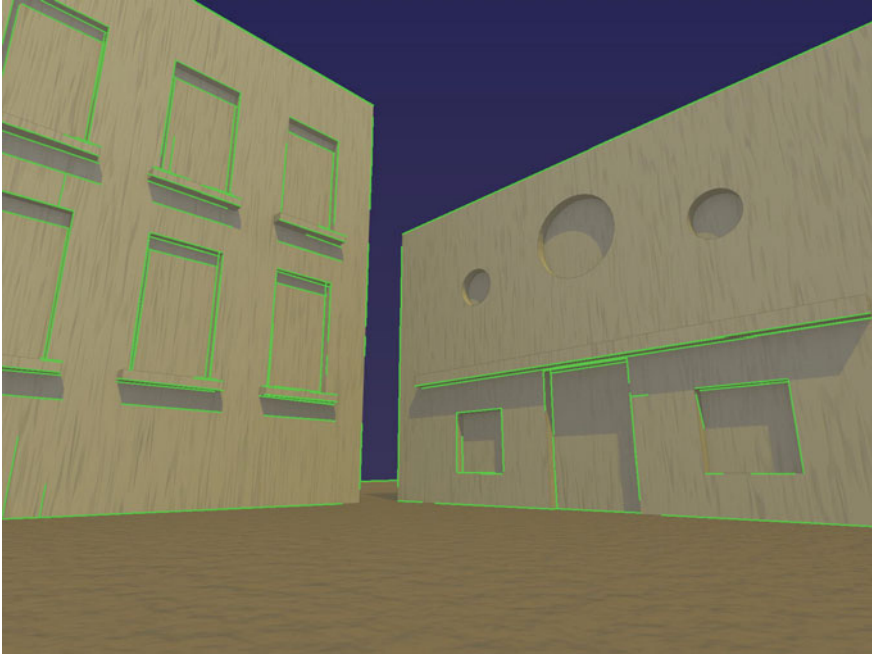


Fig. 11 After the application of the Hough transform, we obtain the straight lines in the image

maximums in the voting matrix determining the parameters of the detected lines. In Fig. 11, we show the detected straight lines after applying the Hough transform on our synthetic render.

4.4 Straight Line Classification and Optimization of the Camera Angles

As a final step before applying the optimization algorithm to detect the orientation of the camera and to rectify the image, we need to sort the lines into different groups: The vertical and horizontal lines in the façade that we want to rectify and lines corresponding to other façades or elements in the image. To do so, we propose the use of a simple clustering algorithm that takes into account the θ angle of the lines (with respect to the bottom side of the image). We can also compute the vanishing points of the clustered lines (which are supposed to be parallel in the real world if they are horizontal and vertical lines in the façade in order to detect outliers that do not belong to the clusters we need).

In Fig. 12, we can see the different clusters of lines that we can obtain: In green, we have the horizontal lines in the façade of interest. In blue, we have marked the

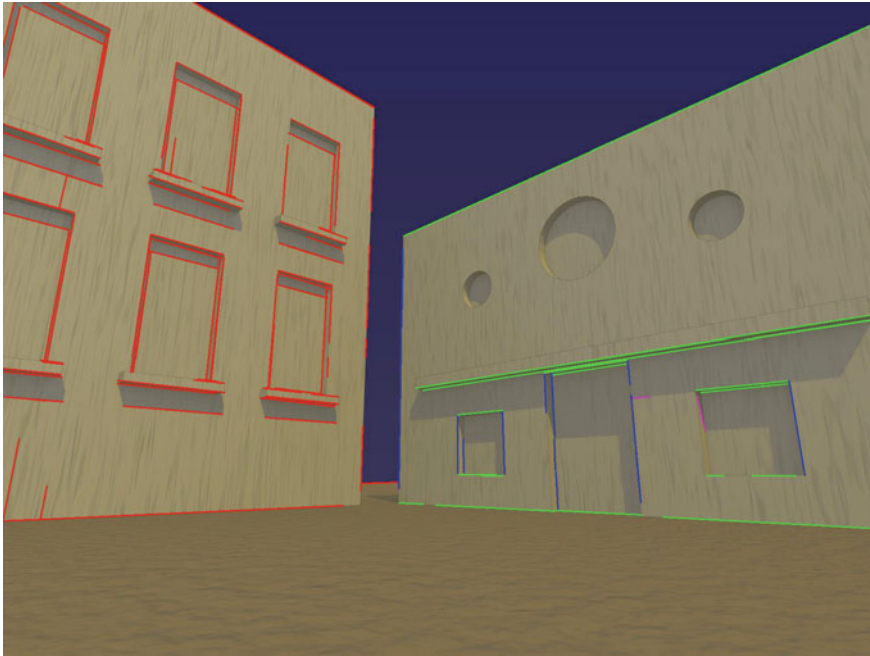


Fig. 12 After clustering the lines, they are classified into the horizontal (green) and vertical (blue) lines in the façade that we want to rectify. Other lines correspond to the other façade (red) or even to different features in the image as shadows or textures (purple)

vertical lines for the same façade. In red, we have the lines that correspond to a different façade and finally in purple, we have some lines in the façade of interest that correspond to other features in the image (shadows in this case).

At this point, we can run the optimization algorithm with the selected vertical and horizontal lines in order to rectify the image. For the test image, the optimization algorithm gives the values $yaw = 44.9^\circ$, $pitch = 15^\circ$ and $roll = 0^\circ$ which are very close to the values that we used to generate the initial synthetic image. Finally, we can apply the appropriate projective transformation to get the rectified image. In Fig. 13, we show the result of applying the projective transformation. As it can be appreciated, the façade plane is now parallel to the camera sensor, and therefore the proportions of the elements are maintained and the vertical and horizontal lines in the façade are now parallel.

We want to remark that this mechanism to rectify images can be used as a good example of how optimization algorithms work and to show the kind of results that can be achieved with them. Playing with the input straight lines and the possible optimization parameters (not only yaw, pitch and roll but also focal length, lens distortions and so on) is a very graphical example of the strengths of these algorithms.

A final example of the good results that can be achieved to rectify photographs of façades in an automatic manner is shown in Fig. 14. There, we deal with a real photo-

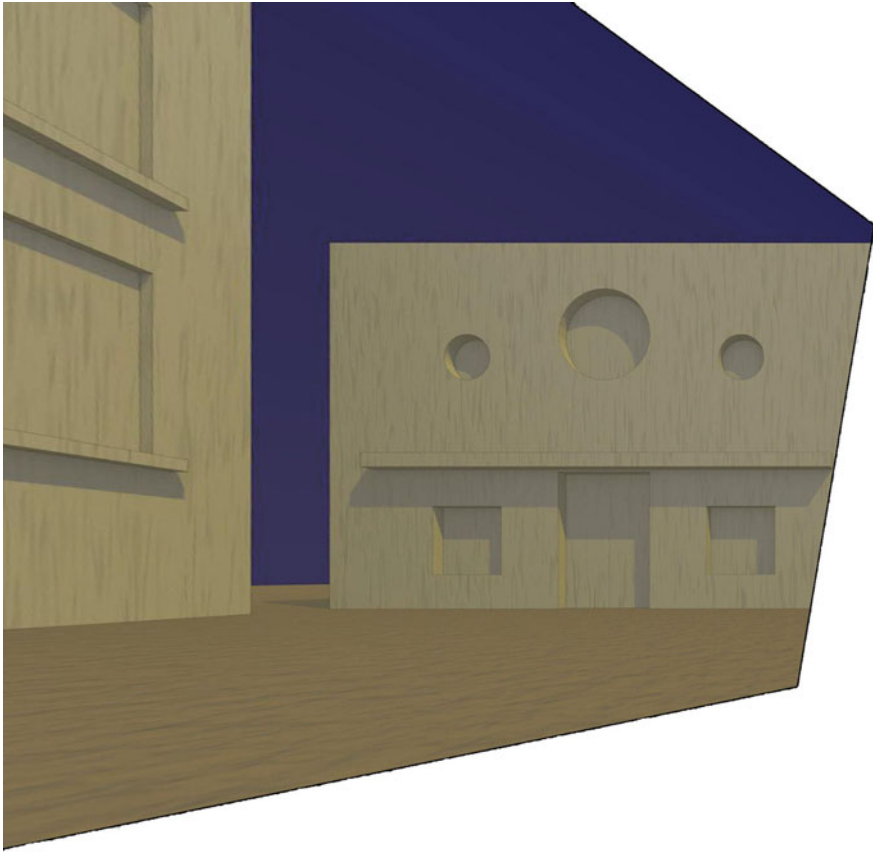


Fig. 13 Rectified image. Now the façade maintains the proportions among its elements and the vertical and horizontal lines are parallel

graph of the façade of the Granada Cathedral. In the same photograph, the façades of the adjoining buildings have been captured. Applying the described techniques, we have been able to reconstruct not only the main cathedral façade but also the other façades in the photograph.

4.5 Constructing an Automatic Geometrical Model for the Façade

Once we have a rectified Image, we can apply different techniques in order to create a geometrical model based on the characteristics of the façade. To do so, we can extend the Hough transform to detect more geometrical features as, for example,



Fig. 14 Real example of the rectification of a photograph of the façade of the Granada Cathedral. **a** Original image, **b** rectified cathedral façade, **c** rectified left façade and **d** rectified right façade

circles, ellipses, squares and rectangles, which are typically used in architecture, and determine their properties (positions, proportions and so on). Moreover, we can also run some pattern recognition algorithms (similar to the ones implemented in MonuMAI) to detect more complex features. As an example in Fig. 15, we have applied an extended Hough transform to detect circles, squares and rectangles. We have successfully detected (in green) the most important features in the façade (windows, door, ledge and round insets). However, the Hough transform has also detected some artifacts due to the effects of the shadows, perspective, textures or even the overall structure of the façade. Please note that the position of each feature and its dimensions and proportions would never be detected with total accuracy: for example, the two square windows may not have been detected with exactly equal size and perfectly aligned.

Once the detection is finished, we'll have a collection of features with their relative positions in the façade and we can study the geometrical relations among them. We can apply different techniques (from AI techniques as neural networks and deep learning to a brute force algorithm) to identify those relations. We would start from some basic relations as "similarity": for example, we could detect that some windows in the façade (rectangles) are similar elements (with approximate sizes and

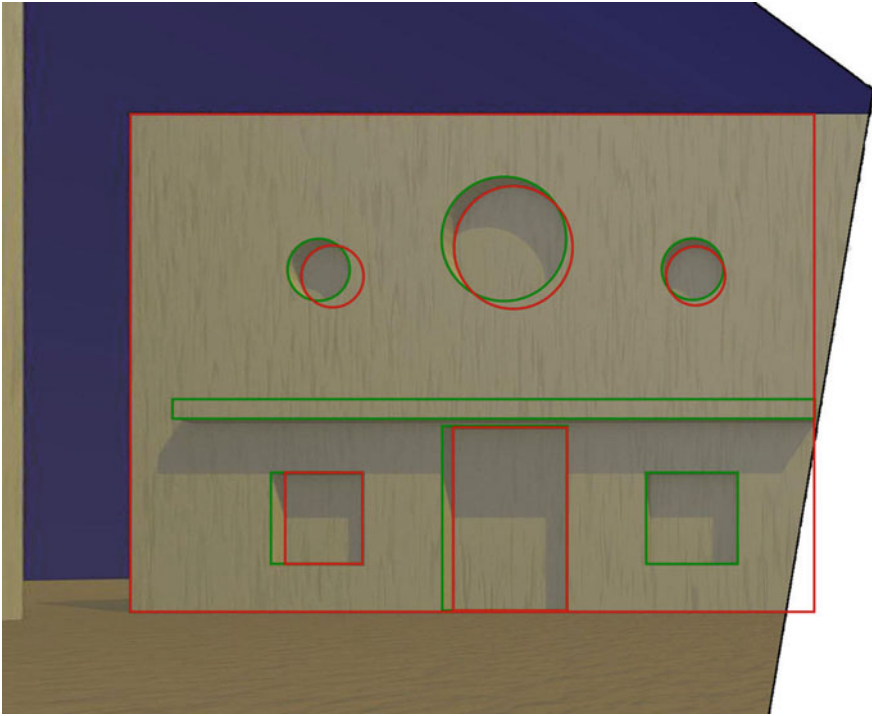


Fig. 15 Application of an extended Hough transform in order to detect circles, squares and rectangles. Green: real detected features in the façade. Red: artifacts detected by effects of perspective, shadows or the overall façade structure

proportions). Later we will try to identify more complex relations as translations, symmetries (along an axis or points) or even more complex geometrical patterns.

In Fig. 16, we show the results of the computation of some geometrical properties for the façade by analyzing the features obtained with the extended Hough transform: In blue, we have detected a vertical axis of symmetry. To do so for each possible vertical axis in the image, we have computed the relative positions of each of the detected features and we have mirrored and matched them (with a certain tolerance for imprecision). We have selected the axis in which the matching of the mirrored features was maximized. In orange and purple, we have depicted some alignment horizontal lines that can be found in the images: the center of the windows and door and the bottom of the circular insets. To obtain those lines, we have followed a similar approach: we have tested all possible horizontal alignment lines in the image against the centers and boundary of the detected features (with some allowed tolerance).

Finally, we want to emphasize that the aims of the proposed method to recreate the geometrical model of façades are multiple:

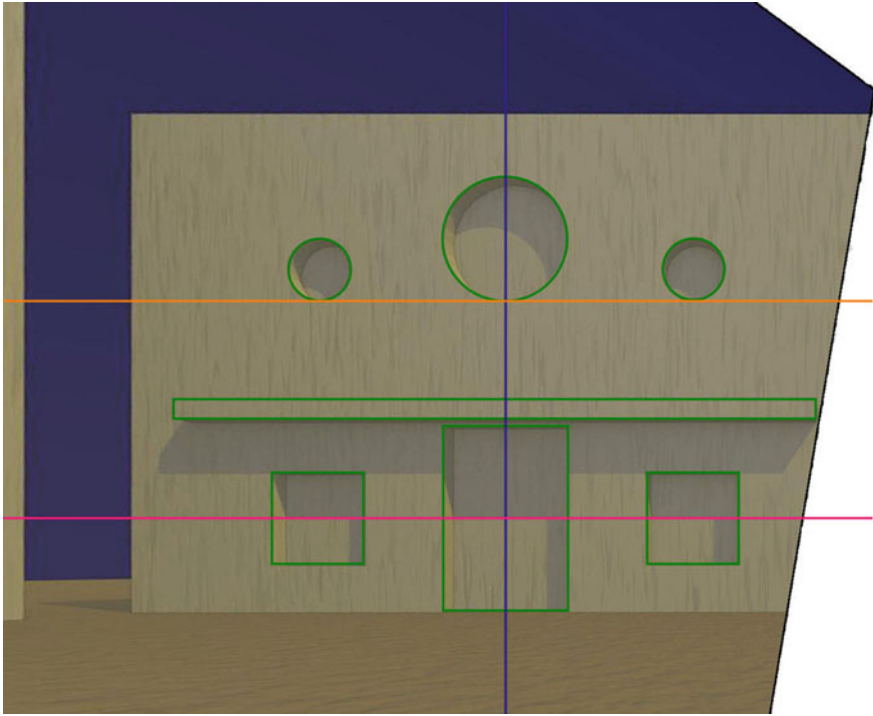


Fig. 16 Detection of some geometrical properties of the façade: a vertical axis of symmetry (blue) and some horizontal alignment lines: center of windows and door (purple) and bottom of the round insets (orange)

- The rectification process is a very good example of how the human brain is very well suited to deal with problems that are difficult to solve by a machine (the effects of the perspective). Therefore, students can better understand how pre-processing the data we work with (in this case an image) is a fundamental step in order to later implement more powerful AI techniques.
- An automated method that “discovers” the geometrical properties of a façade can discharge some of the most tedious work for the person studying it. For example, if the algorithm detects some symmetries among the elements in the façade, the scholars can put all their effort into the interpretation of that geometrical characteristic.
- If the method is applied to a sufficient number of façades, it may be possible to determine which geometrical constructs are used in different architectural styles, thus enriching the mathematical interpretations of each style. For example, presumably in the Renaissance style, we can detect a higher use of the golden ratio than in other styles.

5 Experiences in Education

Our team has built MonuMAI as a tool to be used in education and also in mathematical dissemination. In fact, we have already started teaching it in official maths courses in secondary education as well as in unique activities like science fairs and other educational activities.

In this regard, MonuMAI is part of some Applied Mathematics courses in regular secondary education (15 years). These courses have been carried out, designing a competency scheme for joint use with the subject of Geography and History, with 4 phases:

1. Finding information from buildings in the city of particular historical and cultural interest, highlighting the architectural style of the façade. This affected the *Consciousness and Cultural Expressions* competence.
2. Comparing and verifying the information provided by MonuMAI about the buildings with the information previously collected: *Science and Technology* competence and *Digital* competence.
3. Identifying mathematical ratios on the façade of the visited buildings (MonuMAI Lab): *Mathematical* competence and *Digital* competence.
4. Working in a group in an organized and effective way to complete the global task of gathering and approving the information. *Sense of Initiative and Entrepreneurship* competence and *Learn to Learn* competence.

The activity evaluations for the students were carried out through rubrics, resulting in very satisfactory evaluations in most cases. They were incorporated into the personal evaluation of each student in these subjects.

With respect to the second use in unique educational activities, in addition to lectures, we have done workshops and monumental street walks to show students the possibilities of its application and emphasize the logic with which it reasons, through its use in real cases. Also, we have presented it in various educational and scientific conferences (Descubre Foundation, 2019).

In the lectures and workshops, a set of photographs of especially significant monuments or whose determination leads to cases of interest have been used. Given that almost all students in secondary education own a smartphone and the app is free, this allows each student to do their own experimentation with the software. Among the images provided, there are some with unequivocal identification and others with insufficient identification, for which completion of it is necessary to obtain other shots of the same building (Fig. 17). In these workshops, students have also been encouraged to troll MonuMAI as has been previously mentioned. In the monumental walks, these practices have been carried out live on monuments in different cities, discussing and explaining the outings offered by MonuMAI. Thus, students can see how the probability of determination changes according to light conditions, perspective or distance.

We have also developed a children's teaching kit (MonuMAI_Toy), through the use of simple school games of construction wooden pieces. We have added to Mon-

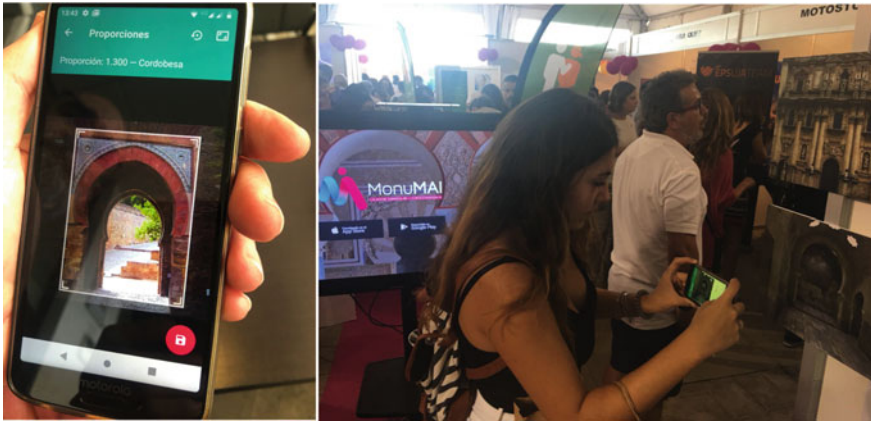


Fig. 17 Use of MonuMAI in education. Left: Example of manual proportions adjustment with MonuMAI Lab. Right: A workshop with some printed images



Fig. 18 MonuMAI for children. Left: recognition of wooden building blocks. Right: Playing and learning with MonuMAI_toy

uMAI the ability to recognize those pieces with their corresponding detected architectural elements (Fig. 18). Children can assemble their wooden building and then apply MonuMAI themselves to detect the assembled elements. The result of the use of this kit has always been a full commitment of the students and children developing a natural curiosity on how MonuMAI carries out such a classification.

Moreover, MonuMAI has also been integrated as a tool within the Educational Innovation Program “Living and Feeling the Heritage” of the Ministry of Education of the Regional Government of Andalucía (Spain) with specified didactic objectives and achieving a high compliance with them, according to the tutors of these projects.

Regarding the application in the education of the AGM, this will be carried out during the present year 2021. So, at the moment of writing this chapter, we still do not have an assessment in this regard.

6 What's Next for the Future?

We have shown some of MonuMAI's strengths and weaknesses in interdisciplinary work with mathematics, AI and art history. MonuMAI is already in its second version, but we hope soon to be able to present the third one with more functionalities and tools, like the previously presented AGM. It is clear that the functionalities provided by the latter will make the recognition of architectural elements and classification of styles much more precise, by being able to rectify the images and make frontal shots of the façades. In addition, the AGM will be possible to be used as a basis for educational work in mathematics on monumental façades.

The algorithmic approach that MonuMAI will use will be twofold: On the one hand, Computer Vision techniques, through the angular use of the Hough transform and its variants, as explained in Sect. 4 and, on the other hand, with the use of Deep Learning through CNN. Both approaches differ greatly, although in our system they complement each other. The AGM will enhance the deep learning algorithm while deep learning will offer us the direction of what kind of mathematical constructs (proportions, geometric elements, etc.) to look for in the monument through the AGM. But what will be the performance of both in purely mathematical tasks?

In some applications, the AGM could clearly be advantageous: for example, to automatically detect proportions in rectangular spans. Since in the model we must already have the segments of the lines that form it, detecting a rectangular proportion only requires making a search in the transformed domain. Two perpendicular segments that join at a point with an added condition: the quotient between their lengths must be the number that marks the proportion. This search, which involves measurement and is carried out directly, seems more efficient and faster to do it through the AGM, than by any other method.

However, other purposes can test both approaches. Like the one proposed in this question: is it possible to automatically learn to detect a symmetry (axial or central) in a façade? And a turn, a translation, or other isometries of the plane?

The most immediate answer is that with the AGM tool, it would be a relatively easy task, at least from a conceptual point of view: Detecting an (axial) symmetry from the AGM would consist of an exhaustive search for possible axes parallel to the vertical or horizontal segments (1 search variable). It also needs a soft verification step for each of these possible axes, of a symmetrical coincidence of the model, or part of it within a precision range. For a translation, the complexity will rise to two variables, always performing the search in the transformed domain. But this conceptual simplicity is tarnished, however, by the possible high computational cost that such tasks would have.

And how would MonuMAI's deep learning approach this task? MonuMAI can be extended to recognize new architectural elements. To do so, we have to expand the dataset with a sufficiently large and representative set of images of the new element to be recognized. As an example, more than 230 images of ogee arches were supplied to be able to recognize this type of arch. However, all these arches share a common geometry: a symmetrical arch that begins with a convex circumferential arc and then passes to another concave arc at an inflection point, joining the two parts at a cuspidal point. Therefore, it is reasonable to think that MonuMAI can be trained with an extended dataset from artificial geometric images to increase the performance of visual recognition. This is a line in which we hope to reach conclusions in the near future.

However, tackling more abstract concepts such as "symmetry" is a clearly much difficult task since it cannot be easily captured in a homogeneous set of images. Each example of symmetry can have a different geometric shape, sometimes so different that it will have very little to do with the other examples, except for the abstract fact of the symmetry itself. Will a deep learning-based AI be able to detect this kind of abstraction? Although we have our own opinions, the answer is, however, still unknown.

In this direction, other possibilities can be the use of more advanced systems based on deep learning or computer vision. These developments may be incorporated in the future into systems focused on teaching mathematics. For example, the creation of applications that are able to detect mathematical objects that are expressed visually. In this group, we can include graphs, lattices, tessellations or even curve characteristics, such as continuity or differentiability. The future presents some exciting possibilities for systems that work with mathematical concepts by means of AI. Systems that could detect properties, and what could be better, offer the possibility of a discussion about the concepts involved. But this will only be useful, and even possible, if today's mathematics education begins a transition from the old paradigms of finalist calculation to the new ones of learning based on logic and the concepts that Artificial Intelligence incorporates.

References

- Ballard, D. (1981). Generalizing the hough transform to detect arbitrary shapes. *Pattern Recognition*, 13(2), 111–122.
- Batanero, C., Chernoff, E. J., Engel, J., Lee, H. S., & Sánchez, E. (2016). *Research on teaching and learning probability*. Springer.
- Batanero, C., Henry, M., & Parzysz, B. (2005). The nature of chance and probability. In G. Jones (Ed.), *Exploring probability in school, mathematics education library* (pp. 15–37). Springer.
- Botana, F., Martínez-Sevilla, A., Kovács, Z., & Recio, T. (2020). *Automatically augmented reality with GeoGebra* (pp. 347–368). Brill, Leiden, The Netherlands. https://doi.org/10.1163/9789004408845_015.
- Canny, J. (1986). A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-8(6), 679–698. <https://doi.org/10.1109/TPAMI.1986.4767851>.

- Chassignol, M., Khoroshavin, A., Klimova, A., & Bilyatdinova, A. (2018). Artificial intelligence trends in education: A narrative overview. *Procedia Computer Science*, 136, 16–24. *7th International Young Scientists Conference on Computational Science*, YSC2018, 02–06 July 2018, Heraklion, Greece.
- Dale, M. R. (2017). Shapes of graphs: Trees to triangles. In *Applying graph theory in ecological research* (pp. 37–53). Cambridge University Press. <https://doi.org/10.1017/9781316105450.003>.
- Dersch, H. (2007). Panorama tools, open source software for immersive imaging. In: *International VR Photography Conference 2007 in Berkeley, Keynote Address*.
- Descubre Foundation. (2019). Monumai citizen science. In *EUSEA Annual Conference of the European Science Engagement Association*. <https://eusea.info/activity/monumai-citizen-science/>.
- Duda, R. O., & Hart, P. E. (1972). Use of the hough transformation to detect lines and curves in pictures. *Communications of the ACM*, 15(1), 11–15. <https://doi.org/10.1145/361237.361242>.
- Fiorucci, M., Khoroshiltseva, M., Pontil, M., Traviglia, A., Del Bue, A., & James, S. (2020). Machine learning for cultural heritage: A survey. *Pattern Recognition Letters*, 133, 102–108. <https://doi.org/10.1016/j.patrec.2020.02.017>.
- Gadanidis, G. (2017). Artificial intelligence, computational thinking, and mathematics education. *International Journal of Information and Learning Technology*, 34(2), 133–139.
- Geogebra web page. Retrieved May 03, 2021, from <https://www.geogebra.org/>.
- Lamas, A., Tabik, S., Cruz, P., Montes, R., Martínez-Sevilla, A., Cruz, T., & Herrera, F. (2021). MonuMAI: Dataset, deep learning pipeline and citizen science based app for monumental heritage taxonomy and classification. *Neurocomputing*, 420, 266–280.
- Martínez-Sevilla, A. (2017a). Artistic heritage meets GeoGebra: A closer view to research on history of art and mathematics dissemination. In *Communication in the Geogebra Global Gathering (GGG-17)*. Johannes Kepler University, Linz (Austria). <https://www.geogebra.org/m/wXrYA38J>.
- Martínez-Sevilla, A. (Ed.) (2017b). Paseos Matemáticos por Granada. Un estudio entre arte, ciencia e historia. Editorial UGR, Granada, Spain.
- Martínez-Sevilla, A. (2020). Matemáticas, tecnología y arte. una propuesta educativa STEAM. *Uno: Revista de didáctica de las matemáticas*, 87, 15–21.
- Martínez-Sevilla, A., & Cruz Cabrera, P. (2021). El módulo constructivo y la orientación del Palacio de Carlos V de Granada: Arte, Geométribolo. *Arte, Individuo y Sociedad (ARIS)*, 33(1), 29–47. <https://doi.org/10.5209/aris.67029>.
- Soycan, A., & Soycan, M. (2019). Perspective correction of building facade images for architectural applications. *Engineering Science and Technology - An International Journal*, 22(3), 697–705. <https://doi.org/10.1016/j.jestch.2018.12.012>.
- Taddia, Y., Gonzalez-Garcia, L., Zambello, E., & Pellegrinelli, A. (2020). Quality assessment of photogrammetric models for facade and building reconstruction using DJI phantom 4 RTK. *Remote Sensing*, 12(19).
- Wojtkowska, M., Kedzierski, M., & Delis, P. (2021). Validation of terrestrial laser scanning and artificial intelligence for measuring deformations of cultural heritage structures. *Measurement*, 167. <https://doi.org/10.1016/j.measurement.2020.108291>.