RECOGNIZING AND EXPLORING AZULEJOS ON HISTORIC BUILDINGS' FACADES BY COMBINING COMPUTER VISION AND GEOLOCATION IN MOBILE AUGMENTED REALITY APPLICATIONS

CARLOS G. R. SANTOS   TIAGO D. O. ARAÚJO   PAULO R. CHAGAS JUNIOR

NELSON C.S. NETO   BIANCHI S. MEIGUINS

Universidade Federal do Pará, Brazil

{carlosresque, tiagoaraujo, prscj, nelsonneto, bianchi}@ufpa.br

Mobile augmented reality (MAR) applications assist users in navigating and exploring their actual surroundings, displaying virtual contents that correspond to objects and scenes in the real world. However, despite the growing popularity of these applications, some experiences can be frustrating when users are unable to correctly recognize Points of Interest (POI), objects, or places they want to visit or obtain more information. The misleading recognition can occur due to imprecise Global Positioning System (GPS) data or a lack of QR codes for interaction. Hence, this article presents a proposal that combines pattern recognition in images with geolocation information to improve the accuracy of the identification of POIs. The usage scenario is the identification of azulejos (tiles) on the facades of historic buildings in the city of Belém of Pará, Brazil. This issue is relevant based on similarities between azulejos and its huge amount of different types, whose variety of designs and colors of geometric forms can make the identification a hard task. The used methods to extract the azulejos' features were the co-occurrence matrix combined with color percentage, and the global positioning data to increase the accuracy of classification because similar azulejos can be geographically far apart. Tests were conducted using six machine learning algorithms (neural network, decision tree, k-nearest neighbors, naive Bayes, random forest, and support vector machine) of different paradigms. The first results show that the pattern recognition in images combined with geolocation information is a promising approach for better identification of the POIs in MAR applications.

Key words: Image Recognition, Machine Learning Algorithms, Mobile Augmented Reality, Azulejos

1   Introduction

The vast majority of mobile augmented reality (MAR) browsers use the sensors of mobile devices to update the users' route or navigational information toward a Point of Interest (POI) they want to visit and obtain more information. However, the GPS (Global Positioning System) used in conventional mobile devices has a low accuracy [1] for situations when the user is near to the POI. This low accuracy causes a poor registration of the virtual markers on the real POI, which frustrates the user when they are unable to find the targeted POI [2].

Another solution is pattern recognition in images to recognize building facades, specific items on the facades of buildings, sculptures, statues, and monuments, among others. The pattern recognition in

images can contribute toward a solution to the limitations found in GPS, either independently or in conjunction with these techniques. Despite the technological evolution of mobile devices, especially regarding processing and memory capacities, adapting and developing image recognition algorithms for MAR browsers remain current challenges.

In this way, this paper presents a proposal for the combined use of image recognition and geolocation information to achieve a more accurate identification of POIs using AR browsers through mobile devices.

A mobile augmented reality (MAR) application is the usage scenario to validate the proposal. The application is called AzulejAR and presents information about historical azulejos (tin-glazed ceramic tilework) in the city of Belém of Pará, Brazil.

There is a large number of azulejos in the city, and there are many similarities between them. Therefore, the possibility of acquiring such information quickly and automatically by taking a picture of a building with an azulejo during a visit is of great importance, especially in a touristic context.

The organization of this article is as follows. Section 2 describes related studies and how they inspired this study's approach. Section 3 explains how the azulejo recognizer was developed, detailing the steps of preprocessing and feature extraction. Section 4 discusses the results of the tests for each adopted strategy. Section 5 presents the final image recognition functionality in AzulejAR. Finally, in Section 6 the concluding remarks and future work are presented.

2   Related Works

MAR applications have experienced a substantial increase in use and are popular in several areas such as medicine [3], games [4], education [5, 6], and tourism [7, 8], among others [9, 10].

One of the important features of AR systems is the insertion of virtual information into the real world to increase the users' perceptions of the reality itself. Computer vision plays a major role in this regard, as it assists in the process of inserting the virtual information into the correct location of the real scene.

MAR register using images requires a careful approach to computing resources, since the mobile device may not always have the maximum processing power for operations, which can affect the user experience. The work in [11] uses image recognition to display augmented information about books on a shelf and maintains the user experience by performing a motion analysis to execute image recognition only when it is suitable.

Searching images efficiently in large databases is also a concern in RAM applications. Pawan [12] describes methods for extracting features with some algorithms, and the SURF proved to be effective for use on compressed datasets in the mobile device. Maidi [13] also uses a large image dataset and proposes that the computation of register runs on a server, which extracts characteristics and performs the search query in parallel with several images.

The mobile visual search is a recurrent and significant problem, mainly due to the advent of mobile devices that come equipped with a camera that serves as a form of rapid capture of the query

image [14]. Also, mobile devices can provide other information that may be relevant to the search, such as user position, sensor data, date and time, and weather conditions.

Visual search can occur only on the mobile device [15] which does not suffer from the network connection and latency problems but does not have scalability for a large database. Thus, it is more common to use a Web server to retrieve information from a client-side query image [16, 17, 18]. Another possible approach is the hybrid visual search that combines the two approaches and selects the fastest and most suitable one at the moment.

Additionally, the image recognition can aid the localization of the user [19, 20] when using a georeferenced image database and a query image to perform the search, thus given the result of image recognition it is possible to obtain the location of the user. Many MAR applications use the user location to perform the register [21, 22, 23, 24].

Some methods can be used to accomplish the visual image search as well as for classification. Bianconi et al. [25] compared image feature extraction methods and machine learning algorithms for the classification of granite tiles. The comparisons made between the methods were performed by using only textures and the combination of colors and textures. Image classification that used the combination of texture features and colors was more efficient than the classification that used only textures.

Yang et al. [26] have shown another evidence of the benefits of combining colors and texture for image classification. Yang et al. [26] developed an interface for blind people to identify clothes' patterns and colors through voice commands and the camera of a mobile device. Their work showed that the combination of color and texture features produces good results.

Among the various existing techniques for texture-based image classification, one that deserves attention is the co-occurrence matrix in grayscale images. Studies show that the co-occurrence matrix is efficient for the proposed task [27, 28, 29].

Therefore, the present work uses the grayscale co-occurrence matrix in the classification of textured azulejos (one or more tiles are repeated along the facade of the building, forming a texture). Additionally, the azulejos have a color pattern that is even richer than the granite tiles used by Bianconi et al. [25], hence the chosen feature extraction uses color characteristics combined with texture pattern in grayscale.

Furthermore, this work presents the combination of image recognition functionality with geolocation information in a MAR application for a more precise identification of azulejos on the facades of historic buildings, without requiring fiducial markers placed on buildings.

3   Proposed Approach

This section describes the strategy adopted to identify a POI using the image recognition of an azulejo. The functionality intends to find a POI using an image of an azulejo. In this case, a tourist could obtain information about a historical azulejo (such as the manufacturer, century of manufacture, country of origin, and other media such as photos and videos) simply by capturing a picture of the azulejo in question.

3.1 Solution Overview

Recognition of the textured patterns of azulejos was developed in stages, primarily to evaluate different implementation strategies. Figure 1 shows an overview of the stages present in the image recognition.

OpenCV [30] was the library chosen for the development of the phases of filter selection, feature detection, and feature extraction. The classification was performed using WEKA [31] data mining software, from the extracted images features.
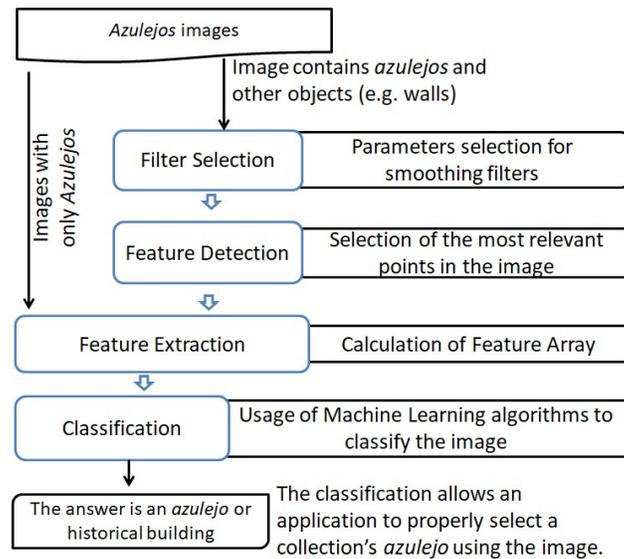
Figure 1. Overview of the stages of textured azulejos recognition.

The feature detection was performed in two ways. The first and most simple method selects equally spaced points in the image to extract image features as a whole. The second method scans the image for corners using the FAST (Features from Accelerated Segment Test) algorithm [32] and uses these corners as entry points to the next stage. Because FAST is intended to find corners, this algorithm can select points of the image excluding smooth surfaces from the extraction of features.

3.2 Setup Database

The image databases contain pictures of historical mansions (open to the public) which have azulejos on their facades and interiors. There were the two approaches for the creation of the image databases used for training recognition classifiers. One image database was created using 191 images captured using three different smartphones, with natural lighting. This image database contains both azulejos and other elements such as walls, windows, and doors, as shown in Figure 2 (A).

The second database created contains 870 images, which are cropped samples of the database with 191 images. This database of cropped images contains only azulejos, as shown in Figure 2 (B).

Figure 2. Two images are representing the two test databases. (A) An image containing azulejos and other elements such as doors and windows. Moreover, (B) an image containing only azulejos.

Since the image recognition functionality will work into an AR tool for mobile devices, some attributes present in these devices can be used to assist in recognition of azulejos. Hence, this article proposes the use of GPS to support the image recognition functionality.

The GPS data were generated from the geographic location (latitude and longitude) of each historic building. The latitude and longitude of each POI were considered to generate more realistic data, and at that point, a random (linear) error of a maximum of 10 meters away from the original position was added. Thus, the machine learning algorithms can determine the azulejo's class type, even when the user's GPS is not accurate.

All the databases generated for training have 12 class types (distinct types of azulejos) to be classified per image. Figure 3 shows four azulejos forming a pattern of each 12 types for illustrative purposes. The photos belonging to the databases had different lighting, positioning, visual noise, and shadow conditions.

Figure 3. The twelve types of azulejos used in the test. These photos are all in good conditions, differently to the photos contained in the test databases, which have different lighting, positioning, and visual noise.

Azulejo recognition was assessed through seven databases with a different strategy for each one. These seven databases were subjected to the same machine learning algorithms, using automatic

selection for the model of each classifier. The bases are listed below, and the term used as a reference throughout the remainder of the article is in boldface.

- **B1**: Base with photos of azulejos and other elements (e.g., windows, doors, wall) and equally spaced points.

- **B2**: Previous (B1) base combined with location (latitude and longitude).

- **B3**: Base with photos of azulejos and other elements and the use of FAST for point selection.

- **B4**: Previous (B3) base combined with location (latitude and longitude).

- **B5**: Base of photos containing only azulejos and equally spaced points.

- **B6**: Previous (B5) base combined with location (latitude and longitude).

- **B7**: Base containing only the location (GPS).

3.3 Filter and Corner Selection

In one of the stages of the construction of the test system, the need to extract image features was evaluated excluding the maximum number of unwanted objects in the image, as in the case of images that contained both azulejos and other elements (e.g., doors, walls, and windows).

Accordingly, it was observed that these unwanted elements have a smoother appearance compared to the azulejos (which have a greater number of edges and corners). Therefore, the FAST algorithm was used to find corners on the images, and two smoothing algorithms were used to prevent noise from being identified erroneously.

The corners identified by the FAST algorithm are passed on to the feature extraction stage. The approach that does not use FAST only selects equally spaced points on the image and moves to the next step. Figure 4 shows the importance of using the FAST algorithm on photos that contain azulejos and other elements.
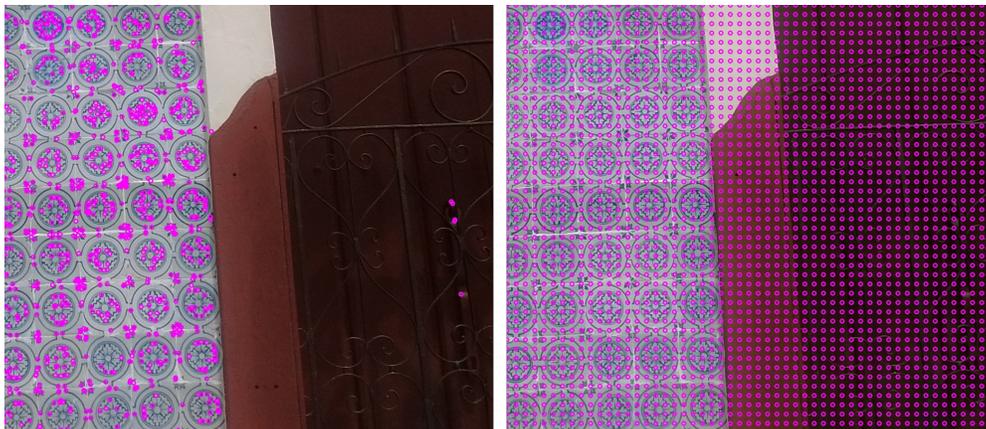

Figure 4. Comparison between points obtained by FAST (left) and equally spaced points (right) in images that contain azulejos and other elements.

Figure 4 (left) reveals that the points from which the features are extracted are primarily concentrated where there are only azulejos, whereas Figure 4 (right) demonstrates the extraction of equally spaced pixels, allowing a large part of what is not an azulejo to pass onto the feature extraction stage.

3.4 Features Extraction

After the selection of filters and corners, the feature extraction of these points (see Figure 4 pink points) begins. This stage works regardless of the strategy adopted in the previous stage; i.e., it does not depend on whether the points were extracted with FAST or whether they are equally spaced points in the image.

The feature extraction allows for the description of images through numeric metrics so that it is convenient for a computer to distinguish these images and visual components. The co-occurrence matrix technique has been used promisingly for classifying and obtaining images containing texture [33, 34, 35], especially when combined with features extracted from the colors in the image [36, 37, 38]. Therefore, the present work uses a combination of these two techniques to classify historical, textured azulejos.

The co-occurrence matrix has several variants, but it is usually applied to grayscale images. The co-occurrence matrix counts how many times two intensities co-occur within the image. In this way, a matrix is created in which the rows are input intensities, the columns are output intensities, and the value inside the matrix corresponds to the number of times that these two intensities have co-occurred.

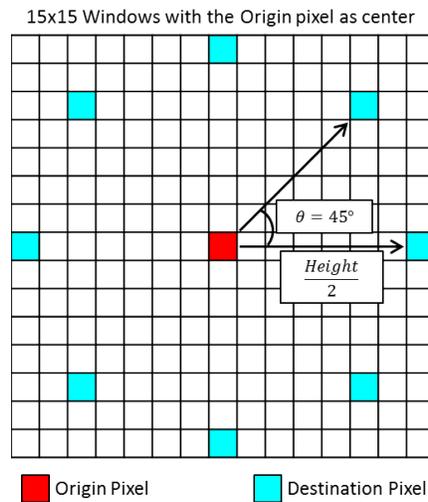### *3.4.1 Co-occurrence Matrix*

Figure 5. Graphical representation of the points recorded in the co-occurrence matrix.

The matrix was constructed from the points selected in the previous stage. The image is converted to grayscale and streamlined to 16 levels of gray instead of 256. For each point chosen in the previous

stage, the level of that pixel in the co-occurrence matrix is recorded in relation to eight pixels at a radius of seven pixels away at equally spaced angles of 45°, as shown in Figure 5. Counting the co-occurrence of a pixel covering 360° makes that metric become invariant to rotation [39, 40].

Then, the following metrics are extracted from that matrix to construct a vector of image features: homogeneity, dissimilarity, energy, correlation, entropy, and contrast [29].

3.4.2 Color Percentiles

To calculate the colors percentage present in the image, points generated in the previous stage were used to select a sequence of windows of the image centered at these points. That is, for every point *(x,y)*, a window of the image was selected with the upper-left corner *(x-(size/2),y-(size/2))* and the lower-right corner *(x+(size/2),y+(size/2))*. The window used in this work has a size of 15x15 pixels.

The image was converted from RGB (red, green, blue; the default of OpenCV) to HSL (hue, saturation, lightness) to calculate the colors percentage. The colors are registered in a vector that contains the following information: percentage of colors (red, yellow, green, cyan, blue, magenta); percentage of 'non-colors' (black, white, gray); and the proportion between colors and non-colors. The colors and non-colors are recorded only for the pixels within the selected windows.

For each pixel, the following verification was made. If the value of *L* is less than 64, then the pixel is black; otherwise, if *L* is greater than 191, then the pixel is white. Conversely, if layer *S* is less than 75, then the pixel is gray. Additionally, the following formula for layer *H* is calculated, where the output *N* represents the colors: *[red=0,yellow=1,green=2,cyan=3,blue= 4,magenta=5]*.

$$N = \frac{H + 15}{30} \mod 6$$

The color vector is concatenated with a vector generated by the co-occurrence matrix. This step returns an image feature vector that will be used by the classification algorithms in the next stage. Then, that stage generates a labeled feature vector for each instance. Twelve different types of azulejos are used, which represents the class types.

3.5 Classification

A file of type ARFF (which is specific for use in the WEKA software), containing the feature vectors extracted from the images, was created for each of the described databases using the OpenCV library.

The machine learning algorithms used in the test are Artificial Neural Network (ANN), decision tree (J48), k-nearest neighbors (KNN), naive Bayes, random forest, and Support Vector Machine (SVM).

For model selection, two classes of WEKA were used: *GridSearch* and *CVParameterSelection*. The *GridSearch* was utilized for the SVM classifier because the defined interval varied in powers of two. For the other classifiers, the *CVParameterSelection* was used, where values varied linearly. In the experiments, the 10-fold cross-validation strategy was utilized for all bases. The varied parameters for each classifier are described in Table 1. Also, the best value found for each parameter are presented.

Table 1. Parameters used in the training stage of machine learning algorithms.

Classifiers	Parameters	Adopted Interval	Best parameters per base						
			B1	B2	B3	B4	B5	B6	B7
ANN	H	[2,33]	10	7	17	6	33	18	31
	L	[0.1,0.9]	0.1	0.1	0.1	0.1	0.9	0.9	0.9
	M	[0.1,0.4]	0.4	0.1	0.4	0.1	0.1	0.4	0.4
J48	C	[0.1,0.5]	0.1	0.1	0.14	0.1	0.14	0.1	0.14
	M	[1,20]	2	3	1	1	2	2	1
KNN	K	[1,30]	1	1	1	1	1	3	3
Random Forest	I	[64,128]	101	64	133	73	122	98	65
SVM	C	$[2^{-5},2^{15}]$	2^9	2^9	2^{15}	2^{15}	2^{11}	2^{10}	2^{15}
	G	$[2^{-15},2^3]$	2^{-7}	2^{-7}	2^{-12}	2^{-12}	2^{-5}	2^{-5}	2^3

4   Experimental Results

The recognizer's objective is to classify the azulejos. The recognition is possible through training the classifiers using the seven bases described in subsection 3.2, each base of the seven (B1 to B7) is divided in a 90:10 ratio for training and test using 10-folds cross-validation, along with the machine learning algorithms and the best parameters obtained in the model selection.

The metric used to evaluate the classifiers was the F-measure. F-measure (also known as F1 measure), by definition, is the harmonic mean of precision and recall. Precision is used as a measure of exactness of the classifier, whereas the recall measures the robustness of a model. Therefore a high F-measure means that the classifier has a good tradeoff between these two measures. As the datasets are unbalanced, analyzing only the accuracy is not enough, since the model can achieve good accuracy classifying correctly only the most numerous classes. Thus, analyzing the F-measure, we evaluate not only the ability to classify some class correctly but also the capacity to discard incorrect classes.

Table 2. Results of the machine learning algorithms related with each database used. Bases from B1 to B4 represent bases that contain azulejos and other elements, B5 and B6 are those that contain only azulejos, and B7 contains only GPS.

Base	*Azulejo* and other elements	Only *Azulejo*	GPS	FAST	Classifier					
					ANN	J48	KNN	Naive Bayes	Random Forest	SVM
B1	✓				**0.921**	0.711	0.787	0.764	0.853	0.762
B2	✓		✓		**0.989**	0.905	0.893	0.958	0.974	0.768
B3	✓			✓	**0.935**	0.727	0.847	0.808	0.850	0.834
B4	✓		✓	✓	**0.984**	0.936	0.916	0.935	0.974	0.834
B5		✓			0.868	0.833	0.869	0.778	**0.896**	0.850
B6		✓	✓		0.965	0.980	0.931	0.964	**0.985**	0.850
B7			✓		**0.957**	0.952	0.952	0.951	0.946	0.950

The results of the training are shown in Table 2, where the rows B1 to B4 represent bases containing azulejos and other elements, the rows B5 and B6 are the databases that contain only azulejos, and the row B7 contains only information from the GPS.

Considering the bases that contain azulejos and other elements, the best results were found for the combination of features extracted from the image and location (GPS). These results were the very close for the two approaches to feature extraction (B2 and B4) using ANN.

For the bases containing only azulejos, the best result was also obtained by combining the features extracted from the images with the geographic location of the user (B6), in addition to using Random Forest. Although the accuracy of using only geographical coordinates provides good results, it is worth mentioning that these tests were performed with a subset of the entire azulejo collection and that, in many cases, there are azulejos in close vicinity of each other. This proximity can cause the errors generated by the GPS to become greater when the basis is extended.

Overall, the ANN algorithm presented the best results in most cases and even in the case where it was not the best (B5 and B6), it exhibited a close F-measure, qualifying it as a potential algorithm for use in the production phase of the tool. Also, the results show that the use of GPS combined with image recognition has obtained better accuracy of classification. However, this database has few classes, and they are far apart. Hence GPS has a higher tendency for better results.

The azulejos chosen in this work are scattered throughout the neighborhood with varying distances, and sometimes away from each other. There are several other azulejos next to each other in the vicinity, and that proximity may end up disrupting the result of techniques that use GPS. As shown in Figure 6, the azulejos can be near each other, confirming the need for a robust technique that does not rely so much on the location of the user.

Figure 6. Two real examples of different types of Azulejos near each other.

The use of FAST for feature extraction is essential for this proximity scenario, since it provides more accurate information, regarding techniques that do not use it.

To confirm the use of FAST as an efficient feature extraction method in this context, we decided to create a proximity scenario with two datasets, B2 and B4, which have azulejos mixed with other elements of the scene. For this scenario, we modified the locations of the POIs at several different scales. We made the approximation by scaling the points considering only part of the total area, in seven new datasets, with the total of eight with the originals datasets, as listed:

- Original: Datasets B2 e B4 with original area;
- -20%: Datasets B2 e B4 with a decrease of 20% of area;
- -40%: Datasets B2 e B4 with a decrease of 40% of area;
- -60%: Datasets B2 e B4 with a decrease of 60% of area;
- -80%: Datasets B2 e B4 with a decrease of 80% of area;
- -90%: Datasets B2 e B4 with a decrease of 90% of area;
- -94%: Datasets B2 e B4 with a decrease of 94% of area;
- -98%: Datasets B2 e B4 with a decrease of 98% of area.

The average distance between the POIs for each dataset used is in Figure 7.
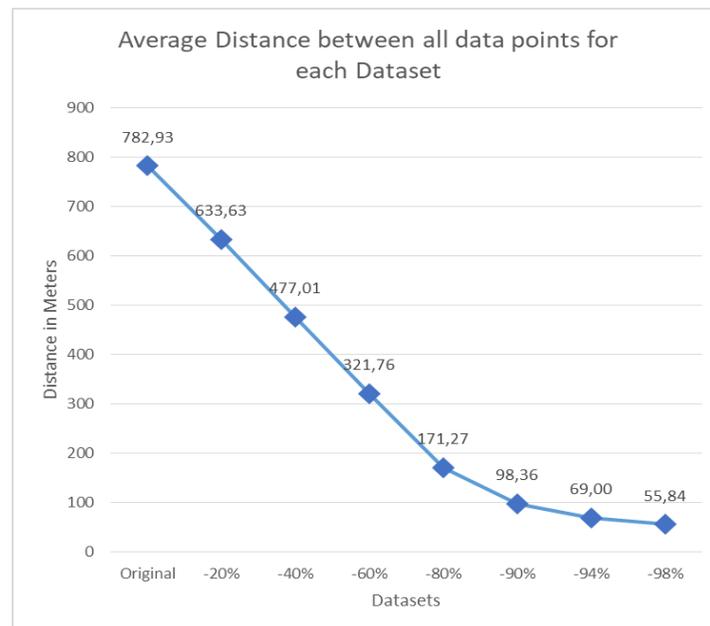
Figure 7. The average distance between all data points for each dataset.

In each base, the POIs were scaled with the same decrease, and with the same approximation used to generate the points of the original datasets, with 10 meters error in the GPS, representing the common error that can be found on the mobile devices. Figure 8 shows the points in each dataset used, with the colors depicting each azulejo.
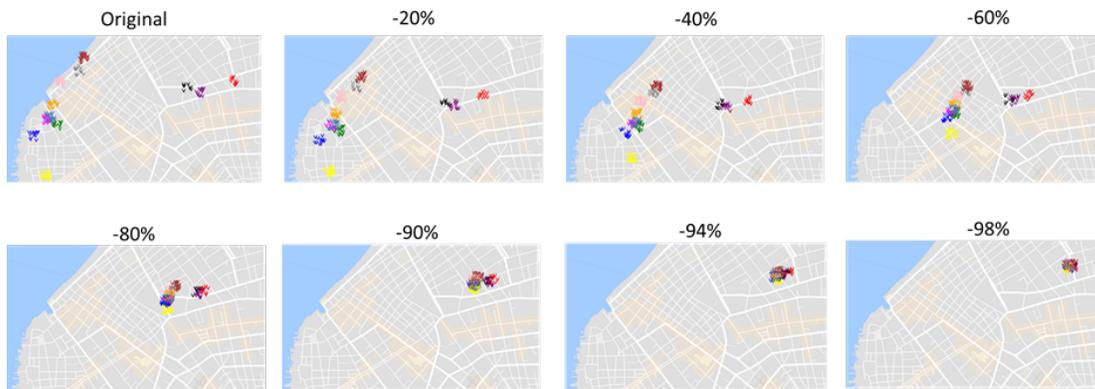
Figure 8. Points of POIs divided by color (each color represents a class), in each decreased area dataset on map. Total area of nearly 1.5 km².

In order to analyze more specifically the improvements that FAST can provide, we considered the two classifiers that obtained the best results in the previous test: Artificial Neural Network and Random Forest. The model selection used the same parameters that were defined in Table 1, but this time considering only the altered and original datasets from B2 and B4, to evaluate how the use of FAST affects the accuracy of the classification as the distance between the POIs are decreasing.

Figure 9 shows the F-measures for each distance condition using ANN. It can be noticed that with a reduction of up to 60% in the distances between POIs, the datasets that did not use FAST showed better results, however with a reduction of at least 80%, the datasets that use the FAST showed better outcomes. This behavior is because the distance between the POIs is decreasing, generating a higher error rate for the GPS. In this context, the usage of FAST tends to overcome this error, so that in situations where POIs are very close, it is not interesting to rely only on GPS.
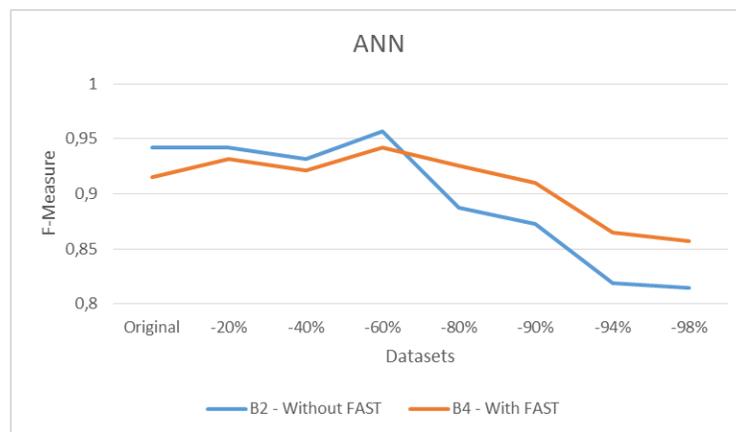

Figure 9. F-measures for each varying the distance condition on databases B2 and B4, using Artificial Neural Network.

Furthermore, the results obtained using Random Forest can be observed in Figure 10, and they also highlight the effectiveness of using FAST for pattern recognition. It can be verified that when using B4

database the results were greater or equal to those using B2 (and B4 was better in most cases). In addition, even with the fall in F-measure (due to the reduction of distances between POIs) the line corresponding to B4 never falls below B2.
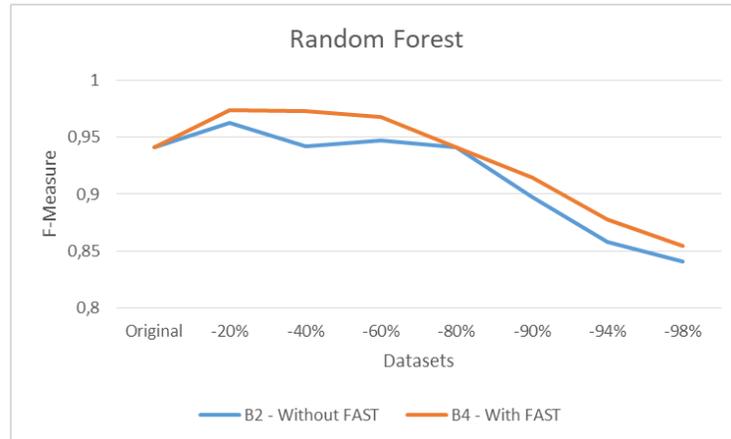

Figure 10. F-measures for each varying the distance condition on databases B2 and B4, using Random Forest.

Considering the results of the two Figures, the approach proposed in this work (B4) proved to be efficient in general, so that it could be verified using two different classifiers. Also, it was possible to validate the proposal in situations where the POIs are near each other, and the GPS error tends to increase.

5   Image Recognition in AzulejAR

The algorithms, techniques, and tests conducted in computer vision and machine learning were all developed on a desktop computer. As soon as the complete process was confirmed to be functional, from image capture to image classification, mobile development was started.

As for the development in tests, the algorithms used in OpenCV with the Python language were translated into OpenCV4Android, an implementation of OpenCV with the required binary files to use the Java library in Android [16]. These algorithms were associated with the AR browser's camera usage in the application.

Regarding the machine learning resources, training models using WEKA were generated. These models can be loaded, and instances can be classified without the need for further training. The WEKA, which is implemented in Java, was used as a dependency for the Android application.

Regarding the interaction, the image recognition functionality was added to AzulejAR, enabling a new form of exploring the content. Thus, the MAR application is functionally able to recognize images to identify POIs and can be used to add virtual information. In the application, the user can perform the recognition of an azulejo by pointing the camera and taking a photo of a building facade containing azulejos. The recognition process is shown in Figure 11.
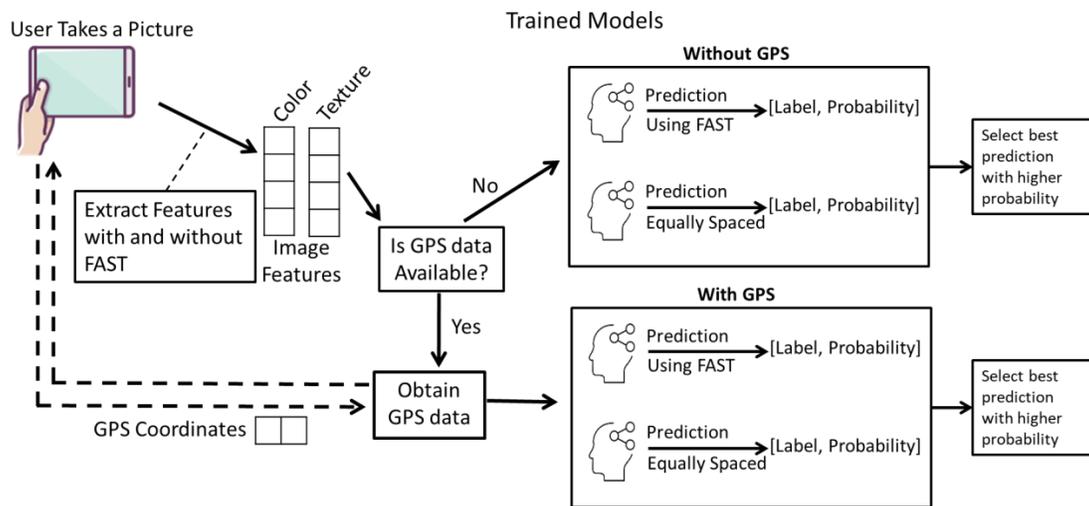
Figure 11. AzulejAR's procedure to perform the azulejos recognition.

With this functionality, it is possible to identify a POI even when the location does not have the required precision for the AR browser to register an icon on the screen accurately. An example is shown in Figure 12 where the register was performed incorrectly. The POIs shown on the screen are far from the user, and the correct POI does not appear even though the user is physically close to its location. In such cases, the image recognizer can circumvent this difficulty of accuracy at close range.
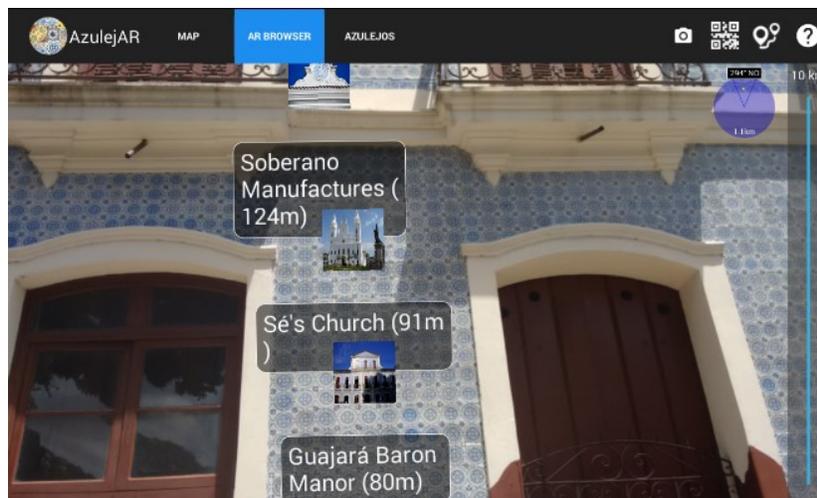

Figure 12. An example of an unsuccessful record due to the proximity to the historic building, muddled by the location.

For recognition, the user takes a photo by tapping the camera icon, and the recognizer begins processing this image, with the extraction of the feature vector. This vector is used as an input to the chosen classification model, and this model returns the likely prediction of the azulejo class type. This result is a list of the possible classes and the percentage of their correctness. If among the sampled

classes (POIs) there is one whose correctness is greater than 90%, then that class is chosen. In case there is more than one class selected by the machine learning model, the application presents the options to the user, allowing them to choose the correct class, as shown in Figure 13.
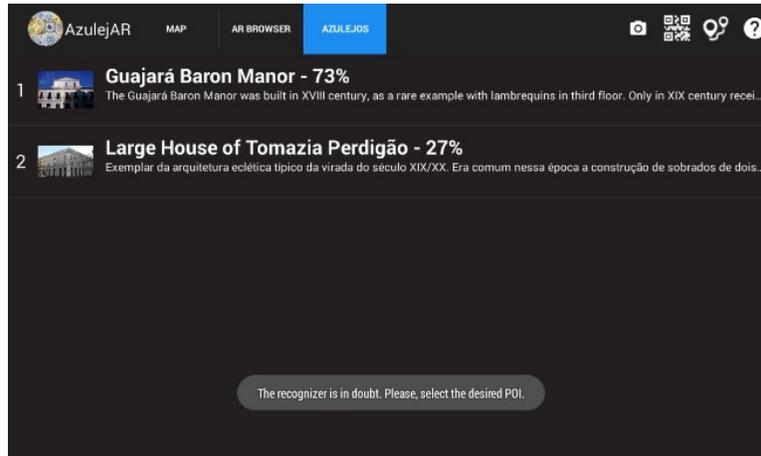

Figure 13. The user can decide whether the POI image is, in fact, the one identified by the machine-learning algorithm.

Thus, even if the recognizer does not have a high certainty rate for the recognition of a POI, the user can make the most appropriate choice, reducing the search scope to only the items selected by the image recognizer.

Finally, when the user touches in a POI from the POI's list, the content about the corresponding POI appears to the user. The Figure 14 shows that user can explore POIs content by text and photos.
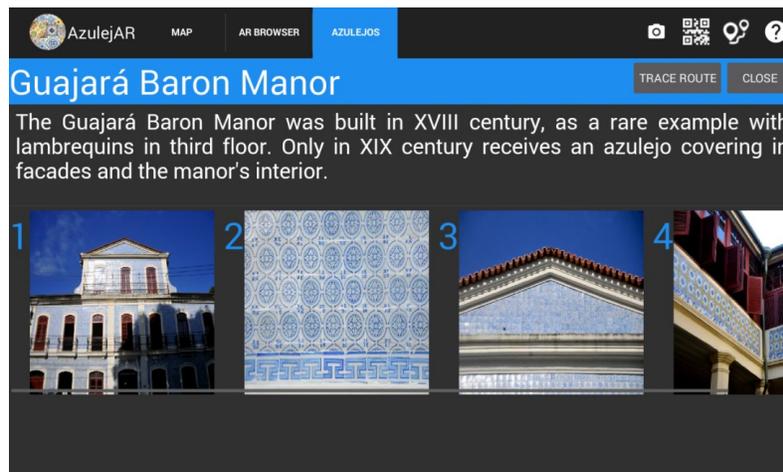

Figure 14. The content is presented to the user after POI's recognition.

6   Conclusion

This work presented an approach for the identification of POIs in a MAR application using image recognition functionality and geolocation data. The usage scenario was an AR application that presents and guides the user through circuits containing historical azulejos. This context proved interesting because of historical azulejos from the city of Belém of Pará; Brazil has a rich diversity of patterns and colors, which makes the task of identifying an azulejo complex. For the accurate recognition of images, different bases, strategies of recognition, and machine learning algorithms were used.

The results showed that it is relevant to combine computer vision and sensor techniques to classify multiple classes of images (which is the case for azulejos). Although using GPS alone provides good results, this precision can decrease with an increase in database size, especially because there are many contiguous buildings (where GPS may fail) with different azulejos (where image recognition is efficient).

These results are positive with respect to a more precise recognition of azulejos and can be a useful tool to retrieve information (stories, texts, websites, images) on the azulejos in an on-site visit. This tool can be provided either embedded in a smartphone or in the form of a web service.

For future work, the focus could be on increasing the database of azulejos to cover the largest possible amount of azulejo types and verifying which machine learning technique performs better when the number of classes is much larger, as there are more than 400 different azulejo types in Belém alone [41].

Tests with users can also be performed to assess usability issues in the application such as where to position the graphic interface elements, how to call the attention of users to this functionality, and how to teach users to use the application for the first time. Other studies can test the performance and retrieval time of each generated model with machine learning techniques and what impact this (delay) has on user experience.

References

1. Kamruzzaman, M., Islam, T., Poddar, S., (2014). Accuracy of handheld GPS comparing with total station in land use survey: A case study in ruet campus. International Journal of Innovation and Applied Studies. Vol. 7, Nº 1, pp. 343–352.
2. Grubert, J., Langlotz, T., Grasset, R., (2011). Augmented reality browser survey. Technical Report, Institute for Computer Graphics and Vision, Graz University of Technology, Austria.
3. Andersen, D, Popescu, V, Cabrera, M, Shanghavi, A, Gomez, G., Marley, S., Wachs, J., (2016). Virtual annotations of the surgical field through an augmented reality transparent display. The Visual Computer Vol. 2, Nº 11, pp. 1–18
4. Liarokapis, F., Macan, L., Malone, G., RebolledoMendez, G., De Freitas, S., (2009). Multimodal augmented reality tangible gaming. The Visual Computer. Vol. 25, Nº 12, pp. 1109–1120.

5. Wu, H., Lee, S., Chang, H., Liang, J., (2013). Current status, opportunities and challenges of augmented reality in education. Computers & Education. Vol. 62, pp. 41–49.

6. Quesada, C., González, D., Alfaro, I., Cueto, E., Huerta, A., Chinesta, F., (2016). Real-time simulation techniques for augmented learning in science and engineering. The Visual Computer. Vol. 32, Nº 11, pp. 1465–1479.

7. Kourouthanassis, P., Boletsis, C., Bardaki, C., Chasanidou, D. (2015). Tourists responses to mobile augmented reality travel guides: The role of emotions on adoption behavior. Pervasive and Mobile Computing Vol. 18, pp. 71–87.

8. Yovcheva, Z., Buhalis, D., Gatzidis, C., (2012). Overview of smartphone augmented reality applications for tourism. e-Review of Tourism Research. Vol. 10, Nº 2, pp. 63–66.

9. Billinghurst, M., Clark, A., Lee, G., (2015). A survey of augmented reality. Foundations and Trends in Human-Computer Interaction. Vol. 8, Nº 3, pp. 73–272.

10. Huang, Z., Hui, P., Peylo, C., Chatzopoulos, D., (2013). Mobile augmented reality survey: A bottom-up approach. Computing Research Repository, pp. 1–35.

11. Chen, D., Tsai, S., Hsu, C. H., Singh, J. P., & Girod, B., (2011). Mobile augmented reality for books on a shelf. In IEEE International Conference on Multimedia and Expo (ICME) (pp. 1-6). IEEE.

12. Baheti, P. K., Swaminathan, A., Chari, M., Diaz, S., & Grzechnik, S., (2011). Information-theoretic database building and querying for mobile augmented reality applications. In 10th IEEE International Symposium on Mixed and Augmented Reality (ISMAR) (pp. 47-53). IEEE.

13. Maidi, M., Preda, M., Lehiani, Y., & Lavric, T., (2014). Vision-based tracking in large image database for real-time mobile augmented reality. In IEEE 16th International Workshop on Multimedia Signal Processing (MMSP) (pp. 1-6). IEEE.

14. Tsai, F. S., Etoh, M., Xie, X., Lee, W. C., & Yang, Q., (2010). Introduction to mobile information retrieval. IEEE Intelligent Systems, 25(1), 11-15.

15. Hartl, A., Schmalstieg, D., & Reitmayr, G., (2014). Client-side mobile visual search. In International Conference on Computer Vision Theory and Applications (VISAPP) (Vol. 3, pp. 125-132). IEEE.

16. Li, D., & Chuah, M. C., (2013). Emovis: An efficient mobile visual search system for landmark recognition. In IEEE Ninth International Conference on Mobile Ad-hoc and Sensor Networks (MSN) (pp. 53-60). IEEE.

17. Kuo, Y. H., & Hsu, W. H., (2017). Dehashing: Server-Side Context-Aware Feature Reconstruction for Mobile Visual Search. IEEE Transactions on Circuits and Systems for Video Technology, 27(1), 139-148.

18. Chen, D. M., & Girod, B., (2015). A hybrid mobile visual search system with compact global signatures. IEEE Transactions on Multimedia, 17(7), 1019-1030.

19. Zhang, J., Hallquist, A., Liang, E., & Zakhor, A., (2011). Location-based image retrieval for urban environments. In 18th IEEE International Conference on Image Processing (ICIP) (pp. 3677-3680). IEEE.

20. Salarian, M., & Ansari, R., (2016). Improved Image Retrieval for Efficient Localization in Urban Areas Using Location Uncertainty Data. In Multimedia (ISM), 2016 IEEE International Symposium on (pp. 181-184). IEEE.

21. Paucher, R., & Turk, M., (2010). Location-based augmented reality on mobile phones. In IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (pp. 9-16). IEEE.

22. Reitmayr, G., & Schmalstieg, D., (2003). Location based applications for mobile augmented reality. In Proceedings of the Fourth Australasian user interface conference on User interfaces 2003-Volume 18 (pp. 65-73). Australian Computer Society, Inc..

23. Alappanavar, P. B., Kurvey, B., Karad, M., & Bhagwatkar, S., (2013). Location Based Augmented Reality. International Journal of Scientific & Engineering Research, 4(5), 566-568.

24. Geiger, P., Schickler, M., Pryss, R., Schobel, J., & Reichert, M., (2014). Location-based mobile augmented reality applications: Challenges, examples, lessons learned.

25. Bianconi, F., González E., Fernández, A., Saetta, A., (2012). Automatic classification of granite tiles through colour and texture features. Expert Systems and Applications. Vol. 39, Nº 12, pp. 212–218.

26. Yang, X., Yuan, S., Tian, Y., (2011). Recognizing clothes patterns for blind people by confidence margin based feature combination. In: ACM International Conference on Multimedia, pp. 1097–1100.

27. Thakare, V., Patil, N., Sonawane, J., (2013). Survey on image texture classification techniques. International Journal of Advancements in Technology. Vol. 4, Nº 1, pp. 97–104.

28. Patel, B., Jaliya, U., Patel, K., (2013). A survey on content based image retrieval. International Journal for Scientific Research & Development. Vol. 1, Nº 10, pp. 71–87.

29. González, E., Fernández, A., Bianconi, F., (2014). General framework for rotation invariant texture classification through co-occurrence of patterns. Journal of Mathematical Imaging and Vision. Vol. 50, Nº 3, pp. 300–313.
30. Bradski, G., (2000). The opencv library. Doctor Dobbs Journal. Vol. 25, Nº 11, pp. 120–126.
31. Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., Witten, I., (2009). The weka data mining software: An update. SIGKDD Explorations. Vol. 11, Nº 1, pp. 300–313.
32. Rosten, E., Porter, R., Drummond, T., (2010). Faster and better: A machine learning approach to corner detection. IEEE Transactions on Pattern Analysis and Machine Intelligence. Vol. 32, pp. 105–119.
33. Capizzi, G., Sciuto, G., Napoli, C., Tramontana, E., Wózniak, M., (2015). Automatic classification of fruit defects based on co-occurrence matrix and neural networks. In: Federated Conference on Computer Science and Information Systems, pp. 5–32.
34. Subrahmanyama, M., Wua, Q., Maheshwarib, R., Balasubramanian, R., (2013). Modified color motif cooccurrence matrix for image indexing and retrieval. Computers & Electrical Engineering. Vol. 39, Nº 3, pp. 762–774.
35. Chang, Y., Zaman, Q., Schumann, A., Percival, D., Esau, T., Ayalew, G., (2012). Development of color co-occurrence matrix based machine vision algorithms for wild blueberry fields. Applied Engineering in Agriculture. Vol. 28, Nº3, pp. 315–323.
36. Huang, W., Daib, J., Wu, Q., (2015). Image retrieval based on weighted block color histogram and texton co-occurrence matrix. In: International Conference on Information Technology and Management Innovation, pp. 848–853.
37. Li, H., Toyoura, M., Shimizu, K., Yang, W., Mao, X., (2016). Retrieval of clothing images based on relevance feedback with focus on collar designs. The Visual Computer. Vol. 32, Nº 10, pp. 1351–1363.
38. Lin, H., Lin, C., (2010). Image retrieval system based on adaptive color histogram and texture features. The Computer Journal. Vol. 54, Nº 7, pp. 1136–1147.
39. Bianconi, F., Fernández, A., González, E., Caride, D., Calviño, A., (2009). "Rotation-invariant colour texture classification through multilayer CCR". Pattern Recognition Letters. pp. 765–773.
40. Davis, L. S., Johns, S. A., Aggarwal, J. K, (1979). Texture analysis using generalized co-occurrence matrices. IEEE Transactions on pattern analysis and machine intelligence v3. pp. 251-259.
41. Alcântara D, Brito S, Sanjad T (2016). Azulejaria em Belém do Pará - Inventário – Arquitetura civil e religiosa – Século XVII ao XX. IPHAN.