Classification of ancient inscription images on the basis of material of the inscriptions

Conference Paper · May 2021

DOI: 10.1109/ICSPC51351.2021.9451641

N. Jayanthi
Delhi Technological University

Indu Sreedevi
Delhi Technological University

Classification of ancient inscription images on the basis of material of the inscriptions

N. Jayanthi
Department of ECE
Delhi Technological University
Delhi, India
njayanthi@dce.ac.in

Tarush Sharma
Department of ECE
Delhi Technological University
Delhi, India
tarushsharma_2k17ec173@dtu.ac.in

Vinay Sharma
Department of ECE
Delhi Technological University
Delhi, India
vinaysharma_2k17ec185@dtu.ac.in

Shubham Tyagi
Department of ECE
Delhi Technological University
Delhi, India
shubham_2k17ec155@dtu.ac.in

S. Indu
Department of ECE
Delhi Technological University
Delhi, India
s.indu@dtu.ac.in

***Abstract*—Machine Learning and AI has allowed us to process images and help us solve vital problems. In this paper, we are classifying inscription images into three image inscription classes namely stone inscriptions, metal inscriptions and palm leaves inscriptions. Due to decaying materials of ancient inscriptions, the classification of such materials becomes challenging. To address this problem, we are using various feature extraction methods like GLCM, KAZE, BRISK to implement texture based feature detection and subsequently classifying them. Both linear and non-linear methods for feature extraction are being used. The paper is concluded by performing classification of images and also making comparisons between all the feature extraction methods in terms of memory required by the algorithm, time scalability and accuracy score of the method. The paper consists of rich information which would be useful in decision making in classification problems.**

***Index Terms*—GLCM, KAZE, BRISK, ancient inscriptions, texture classification**

I. Introduction

One of the integral part in the field of the image processing is the classification based on texture of image.
Images are stored in the form of mathematical data representing height and width of the images with depth of the pixels representing the color of the image.
Most meaningful features for describing images are very much the same features which human being uses in order to interpret the images. Three fundamental pattern element used by humans to interpret colour images are spectral, textural and contextual features. Average variation in tone in various visible and infrared portions of an electromagnetic spectrum is described by spectral features, whereas the spatial distribution in the variations of tone in a band is described by textural features. Contextual features from around the area under observation, provides valuable insights from the pictorial data blocks. If the images being processed are black and white, then tone and the texture of the image are the most important features.

Texture can be described as the fineness, coarseness or the smoothness of an image. It also tells if the image is molled, irregular, lineated etc. Texture is a property possessed by all the surfaces, be it a piece of wood, a grain of sand or a piece of silk. Textures contains in it the information of the structural arrangement of the surfaces and their relationship with the surroundings.

II. Related Works

Texture is a fundamental characteristic used in identifying regions of interest in an image, extracting features from images is a well researched topic and various models are proposed.

GLCM which stands for Gray Level Co-occurrence Matrix,a feature extraction method known for very long was proposed by Haralick.Since it's inception, GLCM has been one of the most important feature extraction methods and finds use in various application of texture analysis.Fourteen different feature were from GLCM were identified and proposed by Haralick, to specify texture [1].Another model which implements feature extraction using GLCM and classification using CBIR(content based image retrieval) is proposed in [2].

Texture classification can be done in many different ways,a dominant neighbourhood structure is proposed in [3], which incorporates the global image features for this purpose.Invariant features of local textures(ILFT) is proposed in [4] which is a texture descriptor algorithm.

S. Leutenegger et al. introduced BRISK method in 2011 [5], which which uses AGAST algorithm to determine the presence of corners, then incorporating FAST corner score to filter the corners, while working simultaneously in the scale space pyramid searching for a maxima.

Pablo Fernandez Alcantarilla et al. intorduced KAZE features, a 2D feature detection algorithm that exploits nonlinear scale spaces. [6].

III. METHODOLOGY

Inscription images are being classifies into three classes which are stone, metal and palm leaves inscription. To achieve this we performed the following steps.

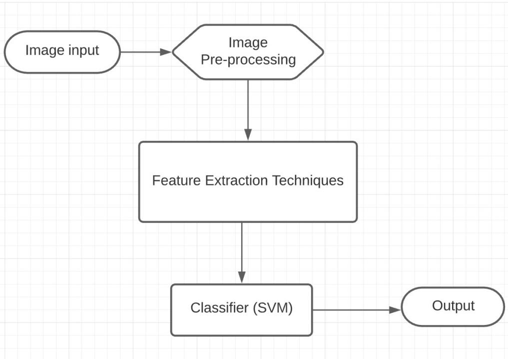

Fig. 1. Flowchart of proposed model

A. *Input & pre-processing*

The data-set consisting of stone, metal and palm-leaves inscription images was prepared from various sources. Due to decaying materials and bad resolution of images classification of data-set is quite difficult. To achieve desired results various pre-processing steps like size reduction of images to appropriate size and in some feature extraction method like GLCM we are converting images into gray-scale. This data is then fed into the various feature extraction algorithms.

B. *Feature Extraction techniques*

We are using various feature extraction methods to extract feature vectors from the images which are essential for the classifier to classifying the images. The feature extraction methods given below like GLCM with its variations, KAZE and BRISK are being used.

1. GLCM

To identify region of interest in an image, texture is a key characteristic. GLCM is used a feature extraction method, where an image can be represented in a matrix that can detect special nature of gray scale pixel distribution in image texture. It evaluates the spatial relationship between pixels, by determining how frequently a pair of pixels with a specific value occur in the image, and thereby creating a matrix.

These are 14 haralick features derived from GLCM matrix:

1. Maximal Correlation Coefficient

$f_1 = (second largest eigenvalue of Q)^{1/2}$
$where,$
$Q(i,j) = \sum_k \frac{p(i,k)p(j,k)}{p_x(i)p_y(k)}$

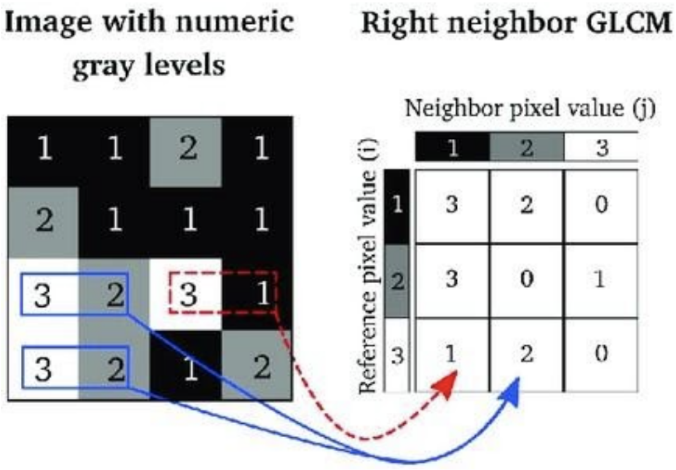

Fig. 2. GLCM matrix

2. Information Measure of Correlation 1

$f_2 = \frac{HXY - HXY1}{max(HX, HY)}$

3. Difference Variance

$f_3 = variance.of.p_{x-y}$

4. Sum Entropy

$f_4 = -\sum_{i=2}^{2N_g} p_{x+y}(i) log\left\{p_{x+y}(i)\right\}$

5. Sum Average

$f_5 = \sum_{i=2}^{2N_g} i p_{x+y}(i)$

6. Sum of Squares: Variance

$f_6 = \sum_i \sum_j (i-\mu)^2 p(i,j)$

7. Contrast

$f_7 = \sum_{n=0}^{N_g-1} n^2 \left\{\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i,j)\right\}$

8. Angular Second Moment

$f_8 = \sum_{i=1} \sum_{j=1} p(i,j)^2$

9. Correlation

$f_9 = \frac{\sum_i \sum_j (ij) p(i,j) - \mu_x \mu_y}{\sigma_x \sigma_y}$

10. Inverse Difference Moment

$f_{10} = \sum_i \sum_j \frac{1}{1+(i-j)^2} p(i,j)$

11. Sum Variance

$f_{11} = \sum_{i=2}^{2N_g} (i - f_4)^2 p_{x+y}(i)$

12. Entropy

$f_{12} = -\sum_i \sum_j p(i,j) log(p(i,j))$

13. Difference Entropy

$f_{13} = -\sum_{i=0}^{N_g - 1} p_{x-y}(i) log\left\{p_{x-y}(i)\right\}$

14. Information Measure of Correlation 2

$f_{14} = (1 - exp[-2.0(HXY2 - HXY)])^{1/2}$

2. KAZE

KAZE feature extraction method is a novel method which works in a non-linear scale on 2d feature detection and description. To identify features in nonlinear scale spaces, non linear diffusion is used which removes the noise and keep all the valid details right in place from the image. One of the simpler cases of diffusion in non-linear manner, variable conductance diffusion is being used. By using Additive Operator Splitting (AOS) schemes, an efficient non-linear space is being built.

A novel mathematical framework called Fast Explicit Diffusion (FED) is used in Accelerated-KAZE Features. Also, a robust Modified-Local Difference Binary (M-LDB) descriptor is computed that uses gradient inputs from non-linear scale space.

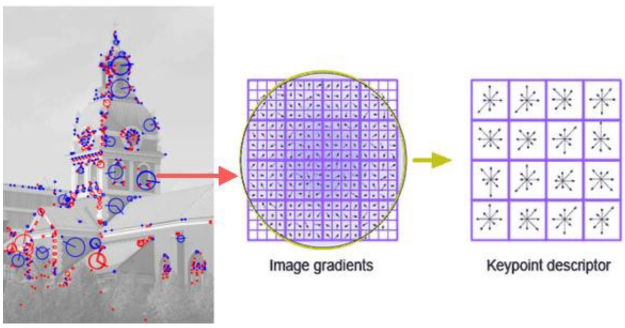

Fig. 3. KAZE

3. BRISK
Three main modules of BRISK algorithm are: descriptor matching, keypoint description and keypoint detection. Initially, the scale space pyramid is being computed, and the extreme points which are stable in continuous scale space of sub-pixel precision are extracted by AGAST (the Adaptive corner detection operator). From the local image neighbourhood, binary feature descriptor of the local image is made ,by incorporating the gray-scale relationship.

Finally, for the purpose of feature matching Hamming distance is used. For the purpose of realisation of scale invariance of descriptors and rotational invariance of descriptors,BRISK algorithm is used which in a multi-scale layer detects feature points and with the help of pixel pairs at along distance, tells the direction of master mode for realising both scale invariance and rotational invariance respectively.

BRISK is a slow method and requires a significant amount of memory for processing.

C. *Classifier*

A supervised machine learning model Support Vector Machine(SVM) was used which is the best fit for group classification problems. By feeding a set of labelled training data to the SVM model, the trained model then was easily able to categorize new input images. The main inclination to use SVM as classifier was because it performs very efficiently on a smaller data-set.

IV. Implementation

A. *Data-set*

We collected a data-set of around 150 images of stone inscriptions, palm leaf inscriptions and metal inscriptions. Data-set was collected from different sources. The images in the data-set were processed to fit our use-case on which we can train and run our model. There are limited algorithms that perform better on small data-sets.

Fig. 4. Stone Inscription

B. *Approach*

We have taken our data-set and trained all the mentioned model to make comparative analysis on all the methods. First we have performed classification using just SVM without using any feature extraction method so we can demonstrate the significance of using feature extraction method. But SVM as a classifier is being used in all the approaches.

1. SVM with kernel = RBF
Here, firstly the images were resized to 256 x 256 pixels and then converted to gray scale and fed to the SVM classifier. The accuracy of the model was found to be 64The accuracy

Fig. 5. Metal Inscription

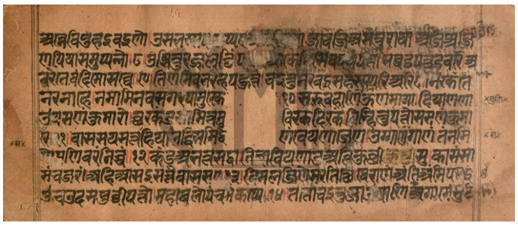

Fig. 6. Palm Inscription

wasn't up to the mark so we decided to look for other approaches.

2. GCLM with gray scale images with 14 features
Here, we fed gray scale images directly into GLCM without resizing them. We decided to use GLCM because it helps in feature extraction and texture based detection. We used haralick texture features which are derived from GLCM matrix and it is a set of 14 different features. For the classification purpose we used an SVM classifier. The output haralick texture features were fed directly to the SVM classifier.

3. GCLM with colored scale images with 14 features
The accuracy of the model was further improved by using coloured images as input instead of gray scale images. Coloured images are 3 dimensional hence 13 GLCM matrices were created which gave better feature extraction property which in turn gave better accuracy, while for gray scale images only 4 GLCM matrices are created per image which resulted in lower accuracy.

4. GCLM with gray scale images with 8 features
Here, we fed gray scale images directly into GLCM without resizing them. We decided to use GLCM because it helps in feature extraction and texture based detection. We used haralick texture features which are derived from GLCM matrix and it is a set of 14 different features. For the classification purpose we used an SVM classifier. The output haralick texture features were fed directly to the SVM classifier. The 8 feature used to train the model were - Sum of Squares: Variance, Difference Entropy, Maximal Correlation Coefficient, Contrast, Difference Variance, Information Measure of Correlation 1, Inverse Difference Moment and Information Measure of Correlation 2.

5. KAZE
The colored images are directly passed as input into KAZE extraction algorithm. There are a number of 2d feature points from which some are selected using a moving window .First, the program uses extracted features to detect key points on the image (the center point of the local pattern). From image to image key point number varies therefore to ensure the resulting feature vector size to remain same we need to add some rules (this is because we can't compare vectors with different dimensions when calculating).Thereafter we compute vector descriptors on the basis of key points. The size of each descriptor is 64. We have 32 such descriptors, so the dimensions of our feature vector is 2048. The output feature vector was directly fed into SVM for classification.

6. BRISK
The colored images are directly passed as input into BRISK extraction algorithm. Here also, similar to KAZE algorithm procedure feature extraction is carried out. BRISK is a very time consuming method. There are a number of 2d feature points used to detect feature in this method, a feature vector of 2048 features is used from which output features are calculated. The output feature vector was directly fed into the SVM for classification.

V. Results & Discussions

A. *SVM with kernel = RBF*

For a smaller data-set just SVM does not give us an appreciable accuracy. The accuracy reaches upto 64.3%.

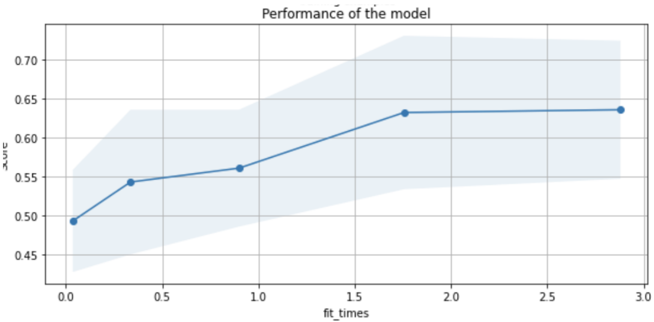

Fig. 7. Performance evaluation for SVM

B. *GLCM with gray scale images with 14 features*

Gray scale images with all 14 features helps in improving the accuracy. The accuracy goes upto 71.2%.

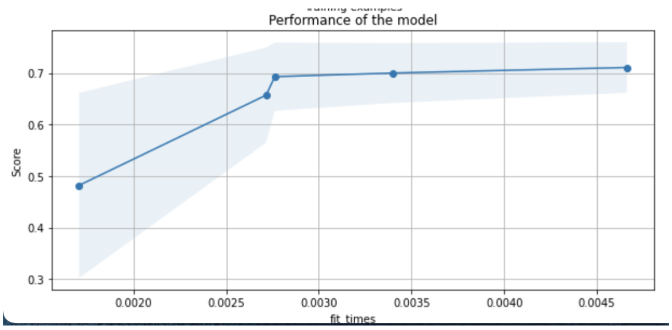

Fig. 8. Performance evaluation for GCLM(gray-scale)

C. GLCM with colored scale images with 14 features

Colored images with all 14 features further improves the accuracy. The accuracy goes up to 73.3%.

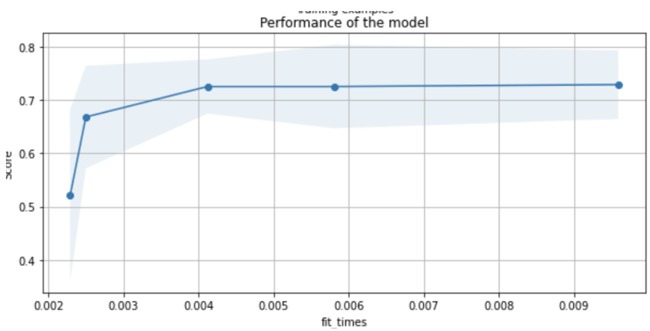

Fig. 9. Performance evaluation for GLCM(colored)

D. GLCM with gray scale images with 8 features

Gray scale images with only 8 features after removing the redundant features from our model. The accuracy goes up to 74.56%.

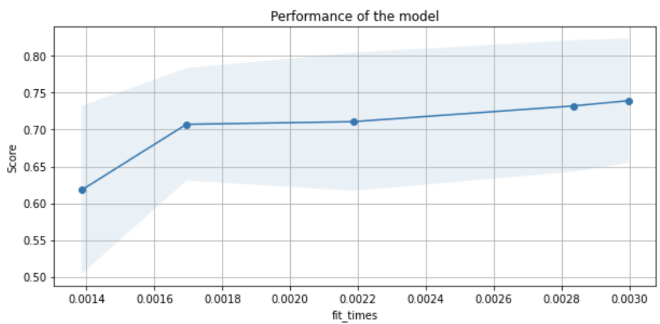

Fig. 10. Performance evaluation for GLCM(selective features)

E. KAZE

KAZE feature extraction method with 1024 feature points was used. The accuracy goes up to 80.5%.

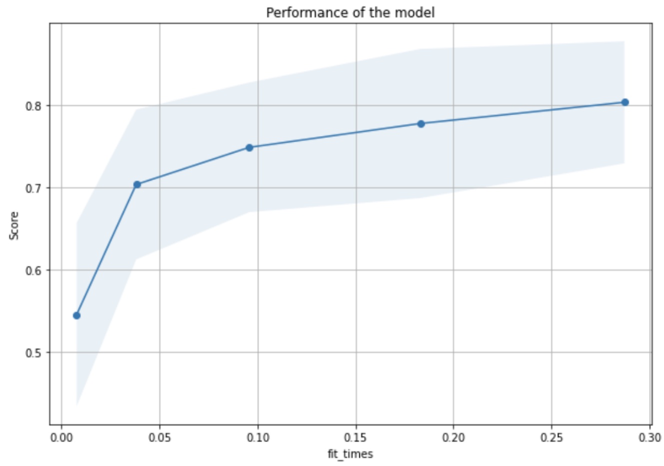

Fig. 11. Performance evaluation for KAZE

F. BRISK

BRISK feature extraction method with 1024 feature points was used. The accuracy goes up to 65%.

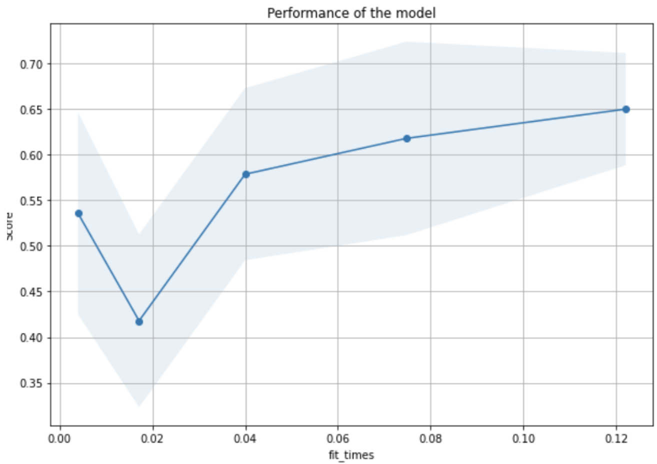

Fig. 12. Performance evaluation for BRISK

TABLE I
ACCURACY SCORES OF EACH APPROACH

Approach	***Accuracy***
SVM(kernel = RBF)	64.3
BRISK (2048 feature points)	65
GLCM(grayscale images & 14 features)	71.2
GLCM(coloured images & 14 features)	73.3
GLCM(grayscale images & 8 features)	74.56
KAZE (2048 feature points)	80.5

Here, exhaustive comparison between GLCM and its variations, BRISK and KAZE is being done. The experimental results presents rich insights which help in making critical conclusions in feature extraction and classification methods.

SVM model being the elementary model has least accuracy. Because it does not consider any features in the image it just classifies in the basis of pixel values.

GLCM considers spatial relationship of pixels hence it was a better fit for classifying the images. A better classification accuracy score was achieved in comparison with SVM. Also, GLCM is computationally most efficient and requires least memory in comparison with other methods. Gray scale images are represented in 2d matrix therefore only 4 GLCM matrices can be obtained for feature extraction but using colored images as input which are represented in 3d matrices from which 13 GLCM matrices can be obtained which assists in better feature extraction hence better classification. Further, there were various features which were having similar values for the three classes of inscriptions (stone, metal and palm leaves) hence, removing the redundant features further improved the accuracy score.

BRISK uses Hamming distance for feature matching which comparatively results in an inefficient feature extraction. Therefore, this results in a poor classification accuracy.

KAZE outperforms all the above feature extraction methods when it comes to classification of images into three classes. KAZE identifies features in a nonlinear scale space using non linear diffusion which removes the noise and keep all the valid details right in place from the image. Therefore, it results in the highest accuracy score.

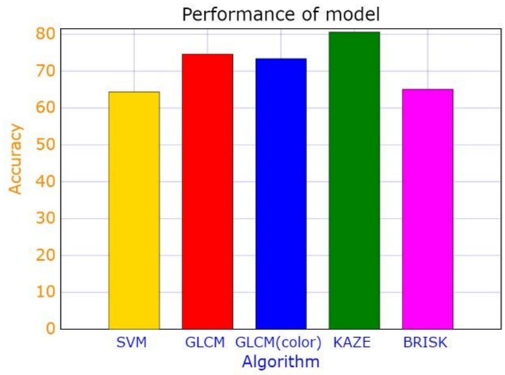

Fig. 13. Performance of model

VI. CONCLUSION & FUTURE WORK

In this paper, we are performing classification of inscription images based on the materials used. Different feature extraction methods are being used to extract feature matrices which are subsequently being used to perform classification. Accuracy of classification increases substantially because of

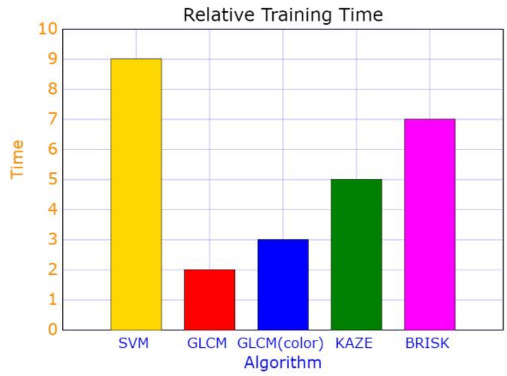

Fig. 14. Relative training time

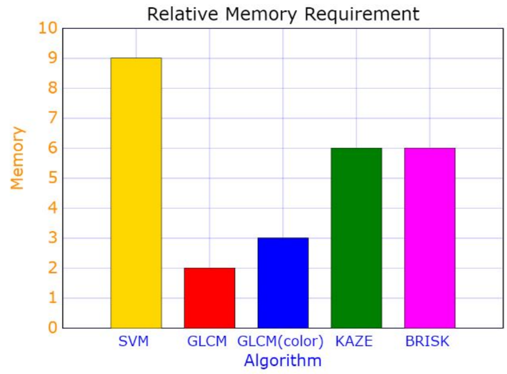

Fig. 15. Relative memory requirement

feature extraction method. Further, we are also evaluating and making comparative analysis between different kinds of feature extraction methods. Comparative analysis shows that KAZE feature algorithm is the most efficient algorithm in terms of accuracy with accuracy score of 80.5%. While KAZE is the most efficient method, taking time scalability and memory requirement into account GLCM turns out to be a better option. Alternatively, in place of feature extraction methods used in this paper neural networks like CNN can be used. But it requires a larger data-set which is one of the challenges and work for future in this field.

REFERENCES

[1] R. Haralick, K. Shanmugam, and I. Dinstein, (1973) "Textural Features for Image Classification", IEEE Trans. on Systems, Man and Cybernetics, SMC–3(6):610–621

[2] Yuslena Sari, Muhammad Alkaff, and Ricardus Anggi Pramunendar, " Classification of coastal and Inland Batik using GLCM and Canberra Distance ",AIP Conference Proceedings 1977, 020045 (2018)

[3] Fakhry M. Khellah, "Texture Classification Using Dominant Neighborhood Structure", IEEE Transactions on Image Processing, 2011, pp 3270-3279

[4] P. Janney and G. Geers, "Texture classification using invariant features of local textures", Image Processing, IET, 2010, pp 158-171.

[5] S. Leutenegger et al., "BRISK: Binary robust invariant scalable keypoints," in IEEE International Conference on Computer Vision, Barcelona, ICCV, 2011, pp. 2548-2555

[6] P. F. Alcantarilla et al., "KAZE features," in European Conference on Computer Vision, Berlin, ECCV, 2012, pp. 214-227