Classifying Architectural Images of Heritage by Using Different Classifiers in WEKA

Eman Qadeer
Department of Computer Science
Riphah International University,
Lahore, Pakistan
emanabid02@gmail.com

Iqra Mushtaq
Department of Computer Science
Riphah International University,
Lahore, Pakistan
iqramushtaq012@gmail.com

Talha Tariq
Department of Computer Science
Riphah International University,
Lahore, Pakistan
talha9828@gmail.com

***Abstract* — In this paper, I implements classification of images using decision tree classifiers in WEKA. The objective of this research is to estimate the performance of decision tree classifiers require small amount of architectural heritage images, and for image features extraction apply 4 filters (1) Fuzzy Opponent Histogram Filter, (2) FCTH Filter (3) Edge Histogram Filter, and (4) Jpeg Coefficient Filter. In this paper, I used classifiers of decision tree involves J48, Hoeffiding Tree, Random Forest, Random Tree and REP Tree. The results show that the Random Forest performs well in classification of architectural heritage images and REP Tree achieved lowest accuracy.**

Keyword: Classification Algorithms, Image Filters, Heritage Images, WEKA.

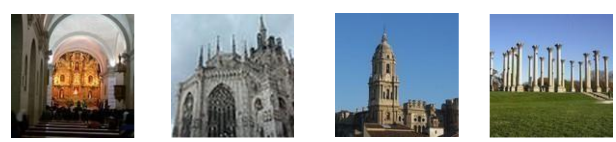
I. INTRODUCTION

Classification of images into dissimilar groups forms on descriptors of images features is image classification. Classification of images is essential task for the architectural heritage, as it allows a finer management of database, and accurate interpretation. Its important to use classifiers as these tasks require a lot of images, may be subject to errors, and for computation require more time. By using decision tree-based algorithms is the one way to perform image classification.

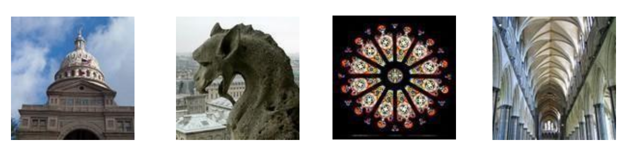
The important part of classification methodology is Decision tree-based algorithms. Main advantage of these algorithms is that data distribution has no assumption, and very fast to calculate. The decision trees classifier are easy to interpret and reliable with image classification, as their constitute consists of a tree with leaves which represent, branches and class labels that use logical conjunction to produce a value build on rule of ”if & then”.

For image classification many algorithms needed additional settings, parameter tuning and adjustments which makes time consuming classification task..

I used a dataset which is publicly available provided for research purposes under Creative Commons license. The images used in this experiment include 480 color images in total (60 for each of eight classes). This dataset contains eight classes: (1) Altar, (2) Apse, (3) Bell tower, (4) Column, (5) Dome(Outer), (6) Gargoyle, (7) Stained Glass, (8) Vault. Before extraction of features and analysis, all images were cropped and resized to the size of 128 x 128.



(a) Altar (b) Apse (c) Bell Tower (d) Column



(e) Dome (f) Gargoyle (g) StainedGlass (h) Vault

Fig. 1 Architectural Heritage Images

The objective of this paper is to estimate how well classification methods work when sustain with image data from a small sample in the architectural heritage surroundings. In particular, four types of features are extracted using the Weka software. Then, build on their values, and applied decision tree classifiers in order to predict the image classes.

The goal of this research is to estimate the ability of decision tree classifiers for architectural heritage images require a small sample.

II. LITERATURE REVIEW

In [4], the authors implements deep learning algorithms to classify images of architectural heritage. For image classification 2 convolution neural networks were applied: (1) Alex Net, and (2) Inception V3. Accuracy results acquire from this research was very good, with over 0.93 mean values. [8]. Authors observed that the deep learning algorithm is very appropriate for classifying images of heritage. And suggested by authors that, if there are limited computational resources, or if the dataset is smaller then use fine-tuning methods [4].

A research by [7] explore the efficiency some data mining techniques like SVM, decision trees and other algorithms to asses land cover changes from remotely sensed data. And It is observed that the decision tree classifier shows good results than other classification techniques.

In [12], an author performs a building facade classification build on an architectural building style, considering the windows. The approach used build on local features learning and clustering [12]. A high classification rate was obtained, proving the efficiency of the proposed method.

In [8], authors implemented an CBIR another dimension in order to assemble images of heritage in 2 classes: 1st class engage activities of human, and the 2nd engage activities of non-human. Authors used a Naive Bayes classifier to classify images build from the edge histogram [8]. In this method, the approach is examined appropriate for classification of heritage images and a fine accuracy was acquired for image classification.

In [15], author uses Weka for classification of cultural heritage images of a small sample using decision tree-based methods. By using Weka Image Filters Package for feature extraction this can be accomplished, and then implements classifiers on extracted features. This research mainly focus cultural heritage and mainly focused on cultural heritage of Eastern Orthodox, such as frescoes and monasteries.

III. RESEARCH QUESTIONS

In [15], Researchers uses three filters of Weka using image filter packages (1) DCT coefficients, (2) edge histogram, and (3) fuzzy color and texture histogram. Then extracted values fed into machine learning algorithms for evaluation of their performances. 4 classifiers were used: (1) Random forest, (2) Hoeffding Tree, (3) j48, and (4) Random Tree. The results observed in terms of accuracy, that Random Forest algorithm gives better accuracy. Hoeffding Tree shows classification accuracy of 0.85. The Random Tree classifier achieved lower accuracy 0.75. The goal of this paper is to find which algorithm is effective and gives better results for any heritage images.

- The first limitation of their work that, They used only 150 images in their research.
- Secondly they use cultural heritage images with just 3 classes.
- There is need to improve the accuracy of algorithm for any heritage images.

IV. METHODS AND MATERIALS

The following steps were applied for architectural heritage image classification involving a small sample.

A. Methodology

The objective of this paper is classification of architectural heritage images using Weka software, on the basis of extracted features. Using the Image Filter package extraction of features is done, a library of java for retrieval of image, by applying different decision tree classifiers for classification of images of architectural heritage, for extraction of image features apply 4 filters (1) Fuzzy Opponent Histogram Filter, (2) FCTH Filter (3) Edge Histogram Filter, and (4) Jpeg Coefficient Filter. In this paper, I used Five decision tree algorithms involve J48, Hoeffiding Tree, Random Forest, Random Tree and REP Tree. For testing the model 10-fold cross validation used. For evaluation results are compared and observed that which algorithm accomplishes best accuracy for classification of images.

B. Data collection

I used a dataset which is publicly available provided for research purposes under Creative Commons license. The images used in this experiment include 480 color images in total (60 for each of eight classes). The dataset included eight classes: (1) Altar, (2) Apse, (3) Bell tower, (4) Column, (5) Dome(Outer), (6) Gargoyle, (7) Stained Glass, (8) Vault. Before the extraction of features and analysis, all images were cropped and resized to the size of 128 x 128. Here are some samples.

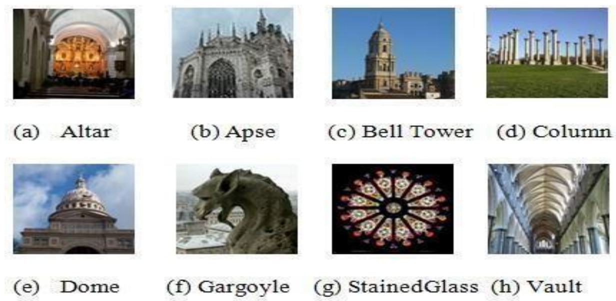
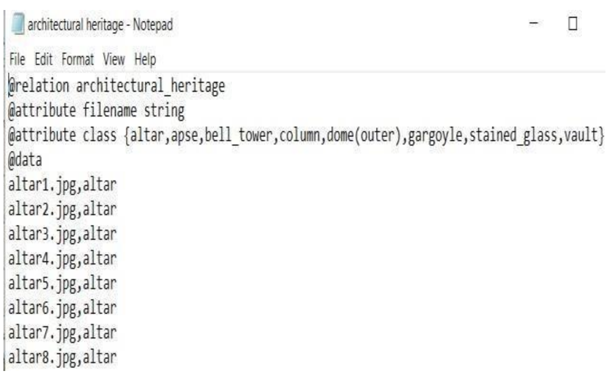


Fig. 2 Dataset

And make the .arff file of images folder, which contains images name and labels.



```
architectural heritage - Notepad
File Edit Format View Help
@relation architectural_heritage
@attribute filename string
@attribute class {altar,apse,bell_tower,column,dome(outer),gargoyle,stained_glass,vault}
@data
altar1.jpg,altar
altar2.jpg,altar
altar3.jpg,altar
altar4.jpg,altar
altar5.jpg,altar
altar6.jpg,altar
altar7.jpg,altar
altar8.jpg,altar
```

Fig. 3 Dataset .arff file

C. *Preprocessing*

Data must be pre- processed by applying filters and data must be transformed into appropriate form of data for exploratory work. So, First step of proposed approach is data preprocessing. When we load the .arff file, Before applying the filter that contains filename and class labels.

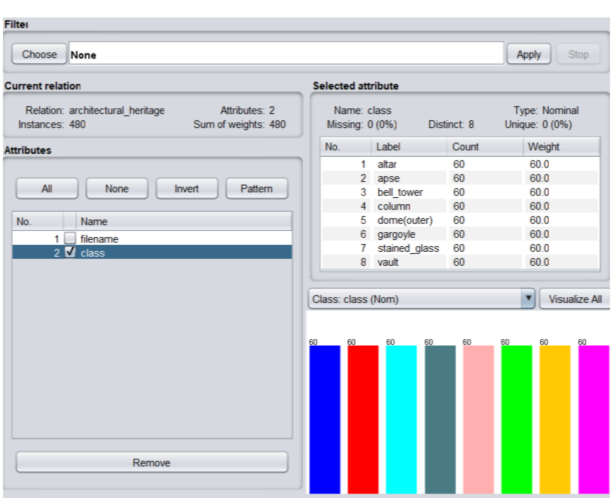


Fig. 4 Dataset Classes

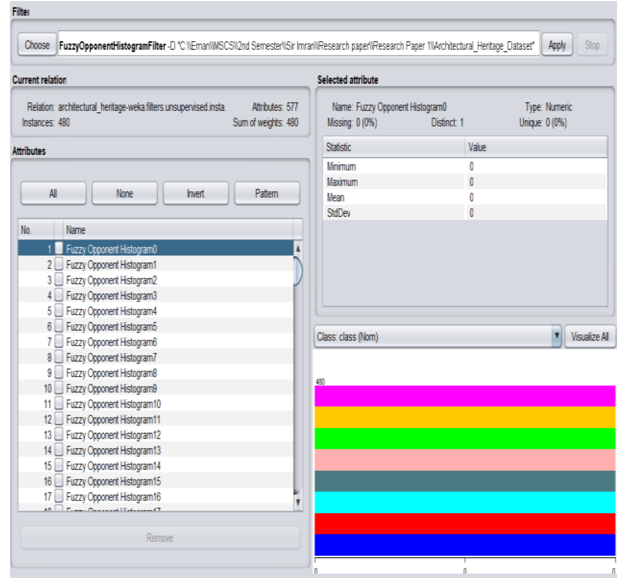


Fig. 5 Fuzzy Opponent Histogram Filter

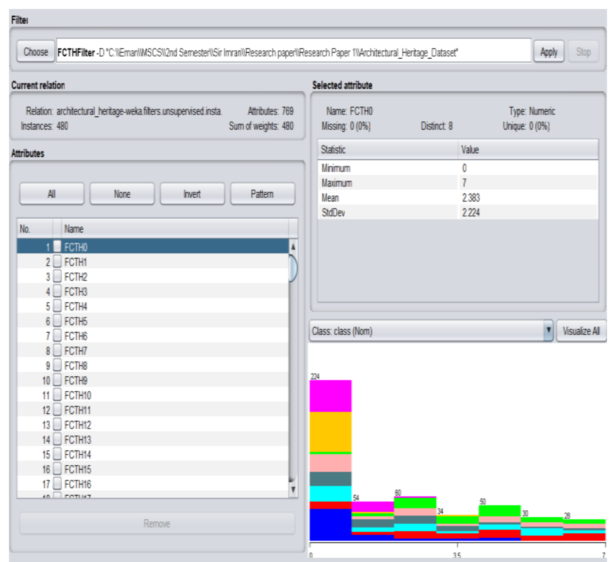


Fig. 6 FCTH Filter

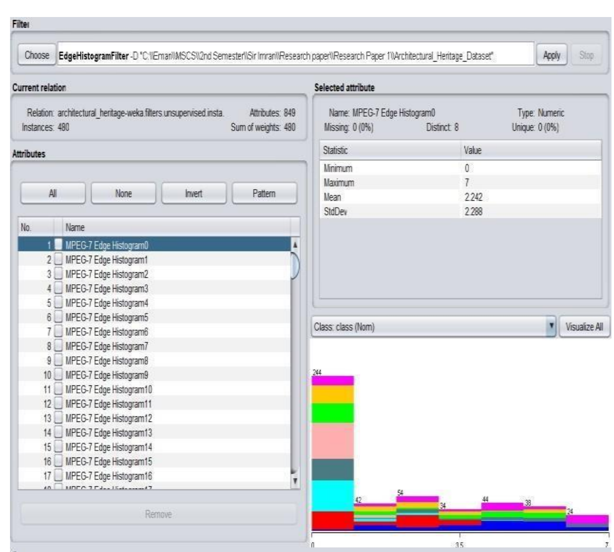


Fig. 7 Edge Histogram Filter

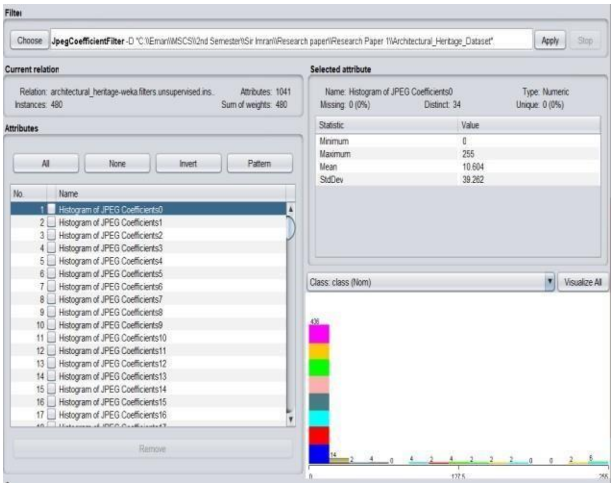


Fig. 8 Jpeg Coefficient Filter

D. Data Classification Process

After applying the Image Filters on architectural heritage image dataset, Several machines learning algorithm applied namely, J48, Hoeffiding Tree, Random Forest, Random Tree and REP Tree using WEKA. After completion of training phase, result obtain from these different algorithms is analyzed and the best performance classifier is selected. And it can be chosen on the basis of different metrics criteria like accuracy, precision, ROC area etc.

V. RESULTS AND DISCUSSIONS

This work uses WEKA data mining tool and used its classification techniques namely Hoeffiding Tree, J48, Random Forest, Random Tree and REP Tree technique applied to classify the data. In this phase of proposed approach all these five classification techniques are applied on the preprocessed image dataset. And the result of each algorithm is calculated and analyzed. Then compare the result of these algorithms with each other.

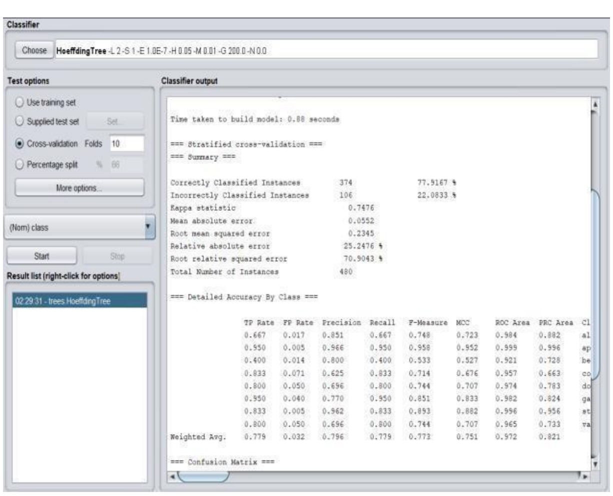


Fig. 9 Hoeffiding Tree

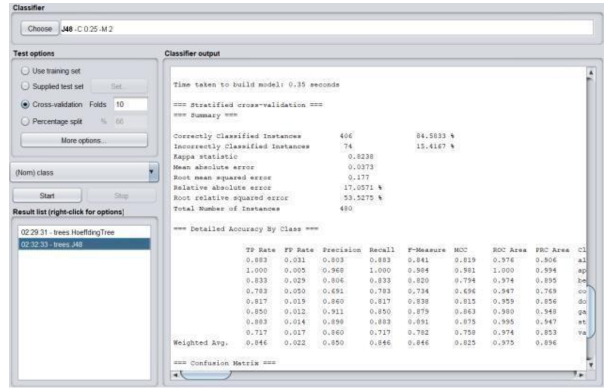


Fig. 10 J48 Classifier

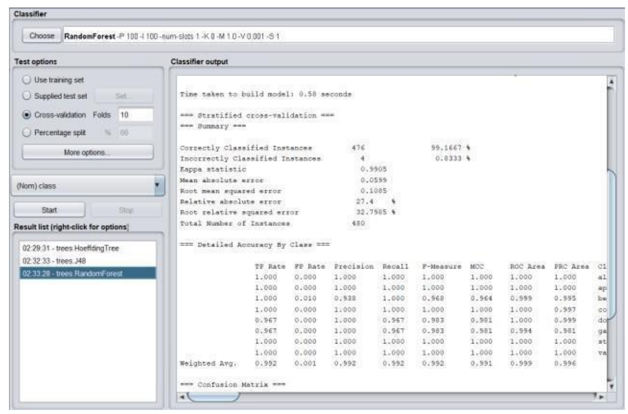


Fig. 11 Random Forest

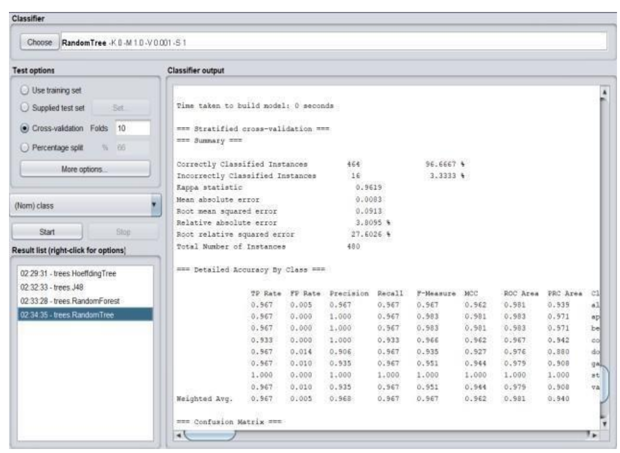


Fig. 12 Random Tree

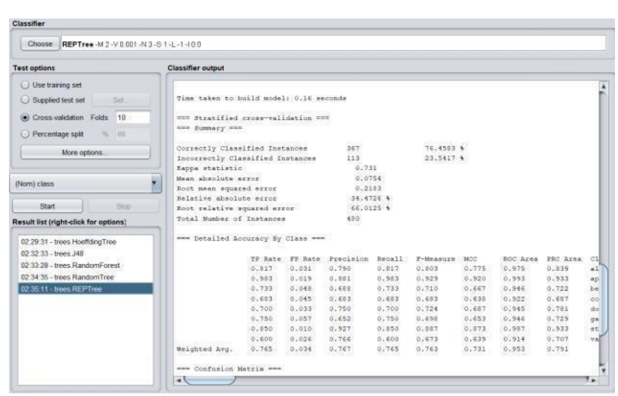


Fig. 13 REP Tree

Table1: Correctly & In Correctly Classifier of Algorithms

Algorithms	Correctly	Incorrectly
Hoeffiding Tree	77.9167%	22.0833%
J48 Classifier	84.5833%	15.4167%
Random Forest	99.1667%	0.8333%
Random Tree	96.6667%	3.3333%
REP Tree	76.4583%	23.5417%

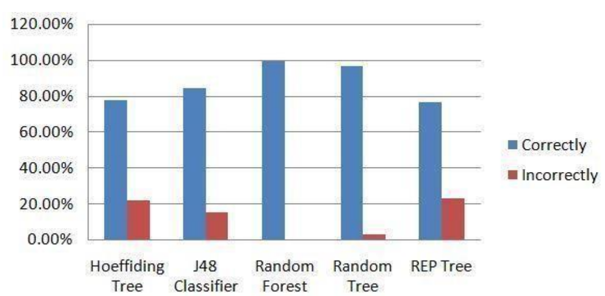


Fig. 14 Graph bar of Correctly & In Correctly Classifier of Algorithms

The above figure 14 shows bar graph of correctly and incorrectly classified of all algorithms. Random Forest has the best accuracy of 99.17% among all other classifiers.

Table2: Algorithms & Accuracy

Algorithms	J48	Hoeffding Tree	Random Forest	Random Tree	REP Tree
Correctly classified instances	84.58%	77.9167%	99.1667%	96.6667%	76.4583%
kappa statistics	0.82	0.74	0.99	0.96	0.73
MAE	0.03	0.05	0.05	0.00	0.07
Precision	0.85	0.79	0.99	0.96	0.76
Recall	0.84	0.77	0.99	0.96	0.76
F-Measure	0.84	0.77	0.99	0.96	0.76
Time taken to build the model(s)	0.35	0.88	0.58	0	0.16

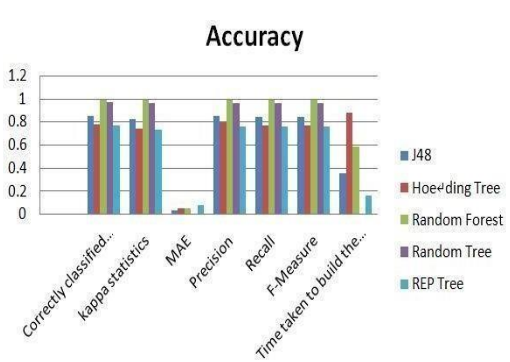


Fig. 15 Accuracy of Obtained Results

The above figure shows that Random Forest has the best accuracy of 99.17% among other classifiers. And the result of these algorithms are analyzed and compared with these decision trees namely (J48, Hoeffiding Tree, Random Forest, Random Tree and REP Tree).

```
=== Confusion Matrix ===

  a  b  c  d  e  f  g  h   <-- classified as
 53  0  0  4  1  1  0  1 |  a = altar
  0 60  0  0  0  0  0  0 |  b = apse
  0  0 50  6  4  0  0  0 |  c = bell_tower
  2  0  2 47  3  2  1  3 |  d = column
  2  0  5  4 49  0  0  0 |  e = dome(outer)
  0  0  5  3  0 51  0  1 |  f = gargoyle
  2  2  0  1  0  0 53  2 |  g = stained_glass
  7  0  0  3  0  2  5 43 |  h = vault
```

Fig. 16 HoeffidingTree Confusion Matrix

```
=== Confusion Matrix ===

  a  b  c  d  e  f  g  h   <-- classified as
 40  0  2  2  0  0  2 14 |  a = altar
  0 57  0  0  0  3  0  0 |  b = apse
  0  0 24 17 15  2  0  2 |  c = bell_tower
  4  0  0 50  6  0  0  0 |  d = column
  0  0  4  6 48  2  0  0 |  e = dome(outer)
  0  0  0  3  0 57  0  0 |  f = gargoyle
  3  0  0  2  0  0 50  5 |  g = stained_glass
  0  2  0  0  0 10  0 48 |  h = vault
```

Fig. 17 J48 Confusion Matrix

```
=== Confusion Matrix ===

  a  b  c  d  e  f  g  h   <-- classified as
 60  0  0  0  0  0  0  0 |  a = altar
  0 60  0  0  0  0  0  0 |  b = apse
  0  0 60  0  0  0  0  0 |  c = bell_tower
  0  0  0 60  0  0  0  0 |  d = column
  0  0  2  0 58  0  0  0 |  e = dome(outer)
  0  0  2  0  0 58  0  0 |  f = gargoyle
  0  0  0  0  0  0 60  0 |  g = stained_glass
  0  0  0  0  0  0  0 60 |  h = vault
```

Fig. 18 Random Forest Confusion Matrix

```
=== Confusion Matrix ===

  a  b  c  d  e  f  g  h   <-- classified as
 58  0  0  0  0  0  0  2 |  a = altar
  0 58  0  0  0  0  0  2 |  b = apse
  0  0 58  0  2  0  0  0 |  c = bell_tower
  0  0  0 56  2  2  0  0 |  d = column
  0  0  0  0 58  2  0  0 |  e = dome(outer)
  0  0  0  0  2 58  0  0 |  f = gargoyle
  0  0  0  0  0  0 60  0 |  g = stained_glass
  2  0  0  0  0  0  0 58 |  h = vault
```

Fig. 19 Random Tree Confusion Matrix

```
=== Confusion Matrix ===

  a  b  c  d  e  f  g  h   <-- classified as
 49  0  2  6  0  1  1  1 |  a = altar
  0 59  0  1  0  0  0  0 |  b = apse
  0  0 44  2  3  7  0  4 |  c = bell_tower
  3  4  4 41  3  4  0  1 |  d = column
  1  0  9  2 42  3  0  3 |  e = dome(outer)
  0  0  5  3  6 45  0  1 |  f = gargoyle
  2  2  0  3  0  1 51  1 |  g = stained_glass
  7  2  0  2  2  8  3 36 |  h = vault
```

Fig. 20 Confusion Matrix REP Tree

VI. CONCLUSION AND FUTURE WORK

A. Conclusion

This paper analyzed the methods in Weka for classification of images using decision tree classifiers. The dataset contains architectural heritage image classification involving a small sample, based on four types of extracted image features: (1) Fuzzy Opponent Histogram Filter, (2) FCTH Filter (3) Edge Histogram Filter, and (4) Jpeg Coefficient Filter. In this paper, I used different classifiers like J48, Hoeffiding Tree, Random Forest, Random Tree and REP Tree. The results indicate that the Random Forest method performs best in classifying a small sample of architectural heritage images, while the REP Tree performs the lowest accuracy. The results shows that, in terms of accuracy, Random Forest algorithm is the best one with accuracy of 99%. The REP Tree classifier achieved accuracy of 76%.

B. Future Scope

In Future work we can increases the size of dataset. And include the comparison of clustering, classification and neural network algorithms for architectural heritage image.

REFERENCES

[1] Random forest in remote sensing: A review of applications and future directions by Belgiu, M., Drgu, L. (2016)

[2] Goyal, A., Mehta, R.: (2012). Performance comparison of Nave Bayes and J48 classification algorithms(2012)

[3] Selection of Best Decision Tree Algorithm for Prediction and Classification of Students Action, Hamoud, A. K. (2016).

[4] Llamas, J., M Lerones, P., Medina, R., Zalama, E., Gmez–Garca–Bermejo, J. Classification of architectural heritage images using deep learning techniques. (2017)

[5] Lux, M., Marques, O.: Visual information retrieval using java and lire (2013)

[6] More, N. K., Dubey, S.: JPEG Picture Compression Using Discrete Cosine Transform (2013)

[7] Land cover change assessment using maximum likelihood classification algorithms and decision trees by Otukei, J. R., Blaschke, T. (2010)

[8] Polpinij, J., Sibunruang, C.: Thai heritage images classification by Nave Bayes image classifier. In Multimedia Technology and its Applications (IDC), (2010)

[9] Prajapati, N., Nandanwar, A. K., Prajapati, G. S.: Edge histogram descriptor, geometric moment and Sobel edge detector combined features based object recognition and retrieval system(2016)

[10] Quinlan, J. R.: C4. 5: programs for machine learning. Elsevier (2014)

[11] Analysis of parametric and non parametric classifiers for classification technique using WEKA. Sahoo, G., Kumar, Y. (2012).

[12] Architectural style classification of building facade windows by Shalunts, G., Haxhimusa, Y., Sablatnig, R. (2011).

[13] Witten, I. H., Frank, E., Hall, M. A., Pal, C. J.: Data Mining, machine learning tools and techniques. Morgan Kaufmann (2016).

[14] Zhao, Y., Zhang, Y., Comparison of decision tree techniques for finding active objects. Advances in Space Research (2008)

[15] Radmila Jankovi´c Mathematical Institute of the Serbian Academy of Sciences and Arts, Serbia rjankovic@mi.sanu.ac.rs (2019).