



STAMP VERIFICATION

Ha Long Duy, Ha Minh Nghia, Phan Duy Hung
FPT University, Hanoi, Vietnam

TABLE OF CONTENTS

1. Introduction
2. Related works
3. Methodology
4. Experiments and Results
5. Conclusion and Future Works

INTRODUCTION



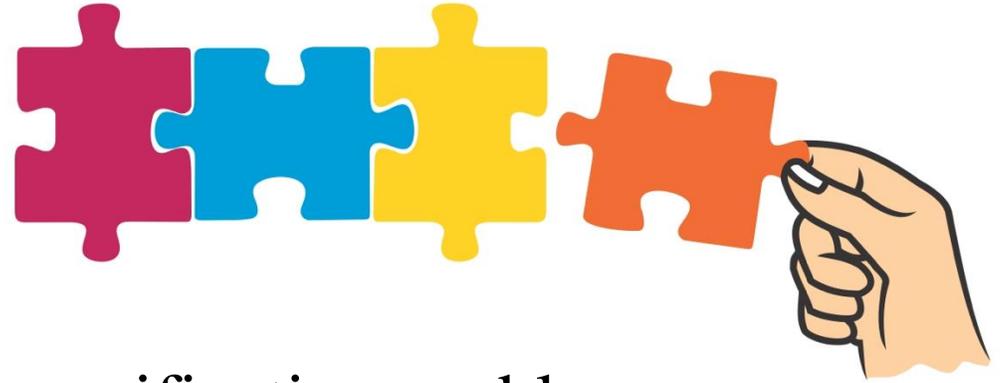
- Every day, thousands of invoices, contracts, certificates and documents are handled. The stamp is one of the most widely used security methods.
- An automated stamp verification method is needed to be developed.

- Stamps vary greatly in shape, size, and color from one company to the next, as well as within a single organization's departments. Stamps come in a variety of shapes, including textual, pictorial, regular (official), and irregular (for fun).
- In this thesis, the process of stamp verification follow 3 stages: Segmentation, classification stamp or non-stamp, classification forged stamp or genuine stamp. Each stage, an simple and efficient approach is developed with various algorithms and models.



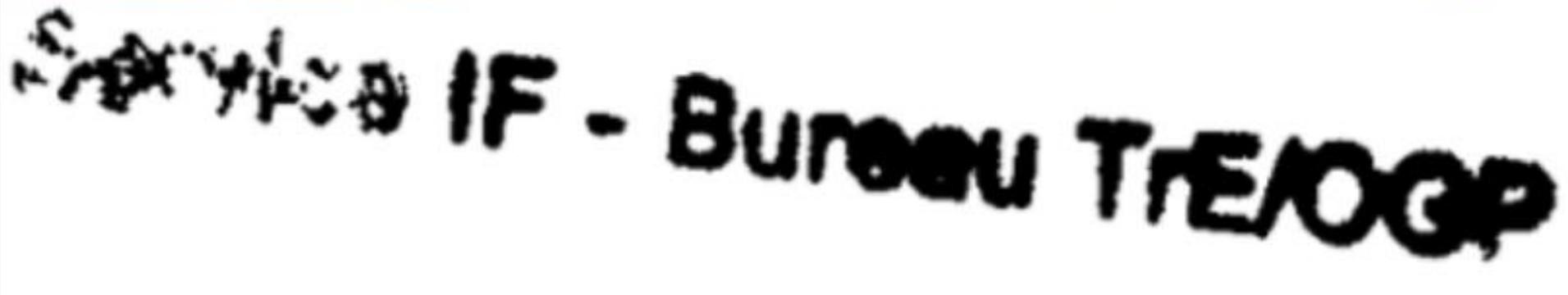
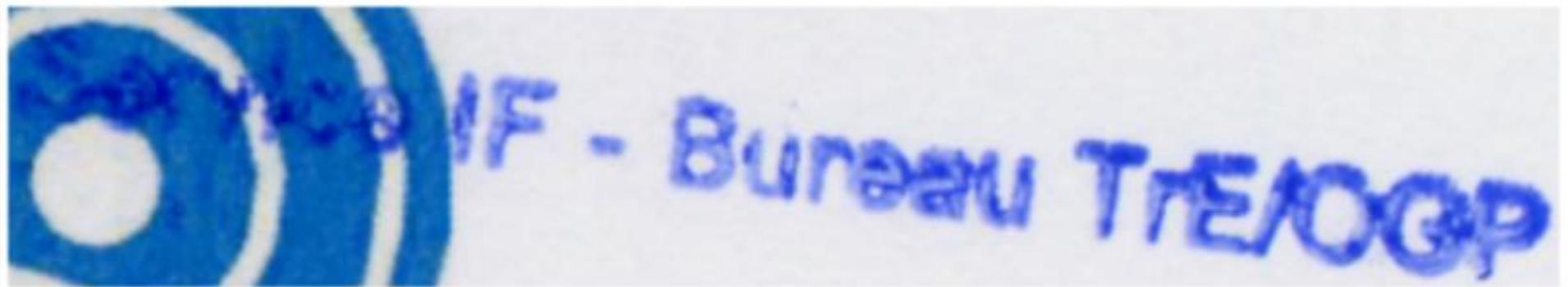
Example of stamps on documents

Contribution



- Develop an efficient approach to stamp verification problem.
- Make the approach easy to implement and efficient to give back good result.

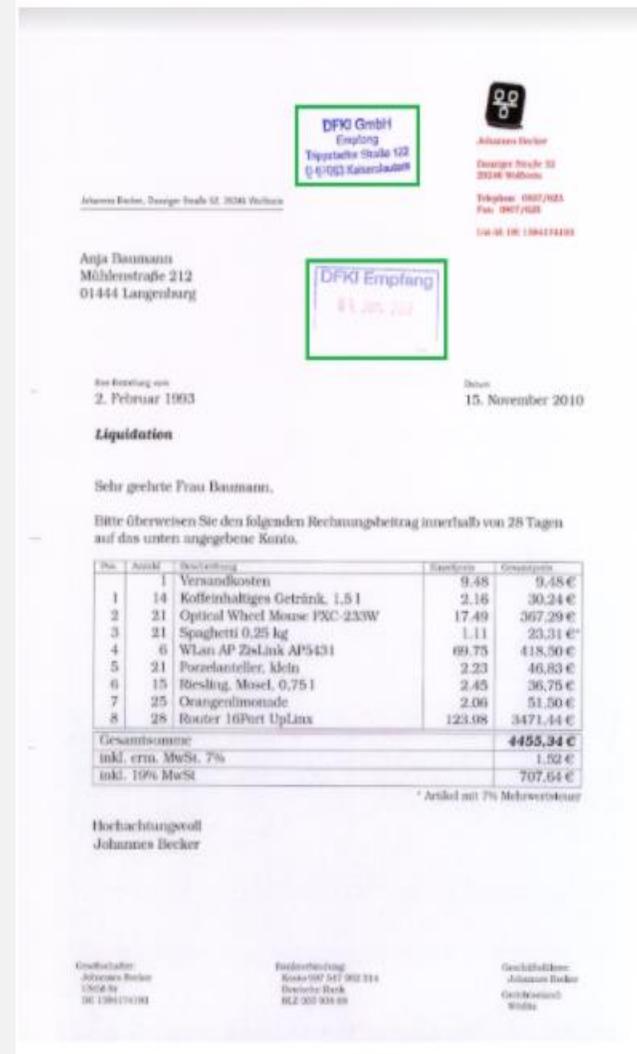
RELATED WORKS



Many previous works only focus on the task of detecting and segmenting stamps from a document. Their ultimate goal is to separate stamps from logos, text, and other information in original document images.

- Other previous works solved the stamp verification entirely, both stamp segmentation and stamp authentication verification.

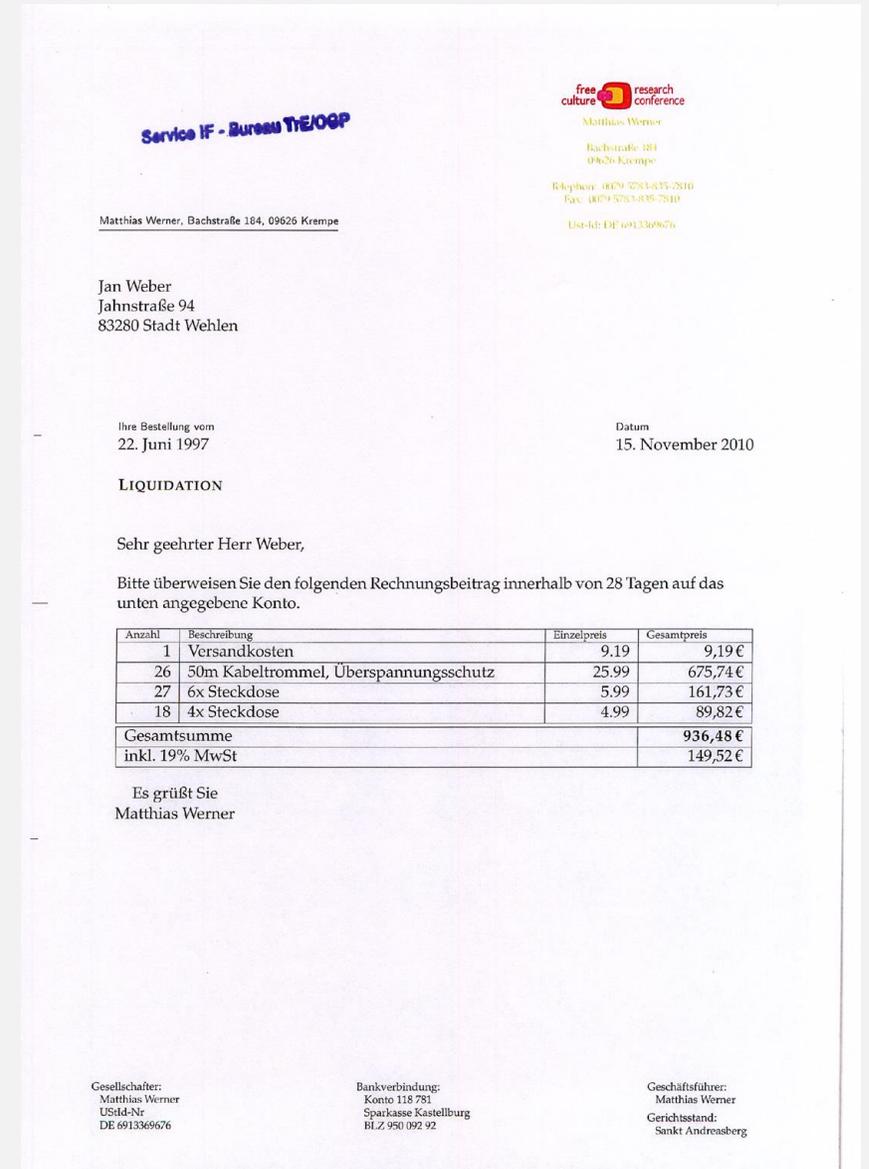
- Finally, 2 works by Micenkov et al developed 3 stages of stamp verification in this paper that we applied in this thesis.



Sample stamps detected marked in green boxes

DATA SET

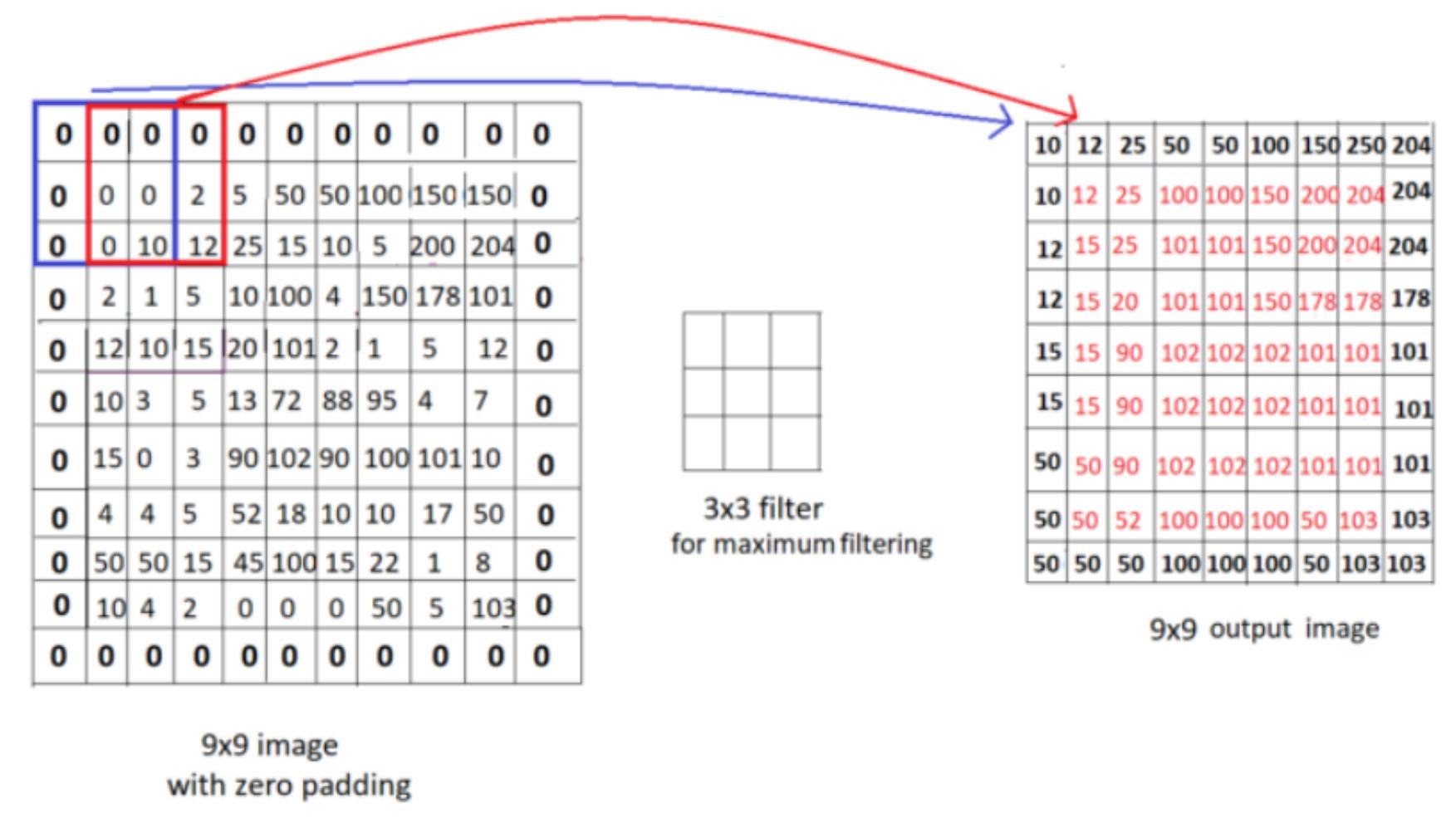
- This thesis evaluated both classification stages, detection and verification, separately with the public data set StaVer containing 400 document pages for evaluation of stamp detection algorithms.
- The reason we choose this dataset is that they are public, free to access, and contain skilled forgeries.



A sample document that contains stamps in the data set

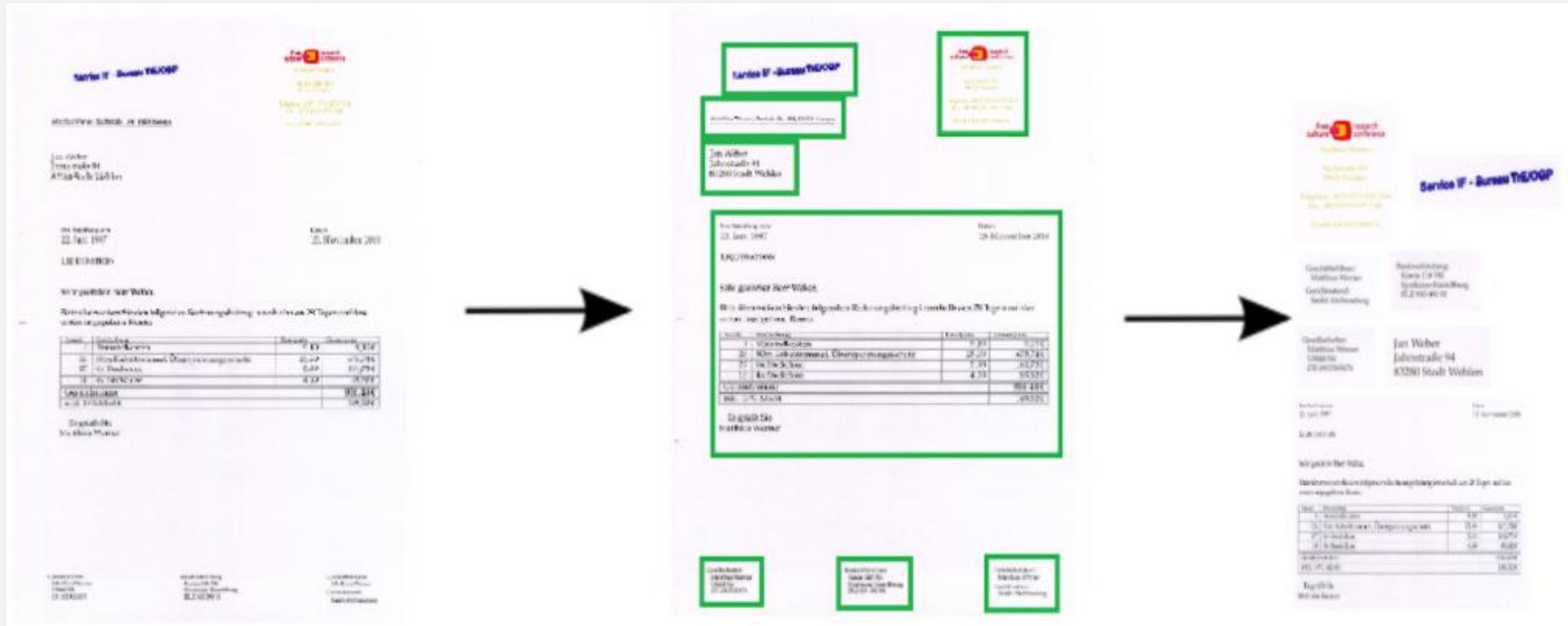
METHODOLOGY

PREPROCESSING



Median filter illustration

SEGMENTATION



The method in the first stage we introduced above has 4 steps: (1) Preprocess image by resize, binarize, bur; (2) Find cluster; (3) For each cluster, detect edges, close figures, find contours; (4) Select largest contours.

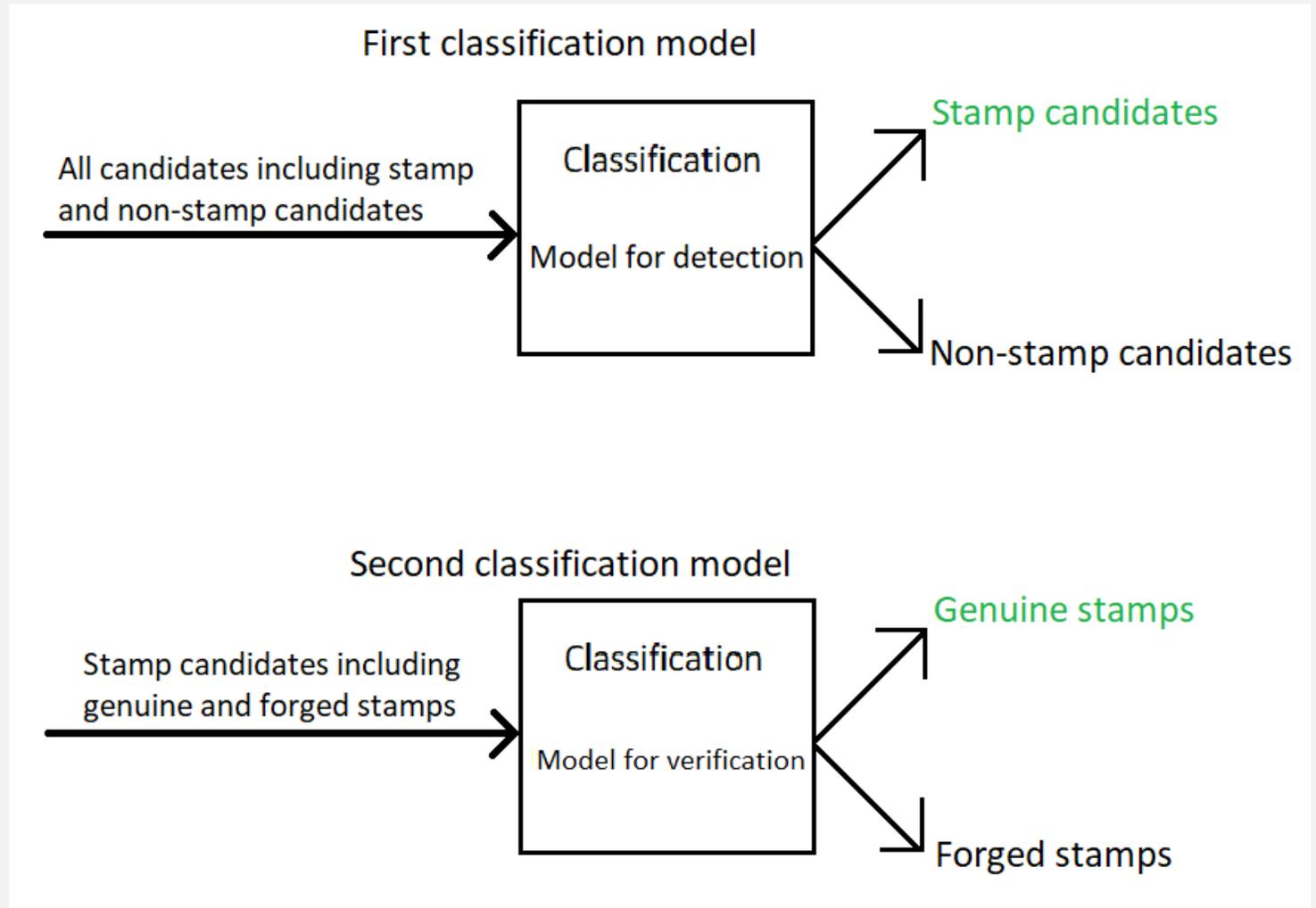


Sample document image with candidates marked in green boxes

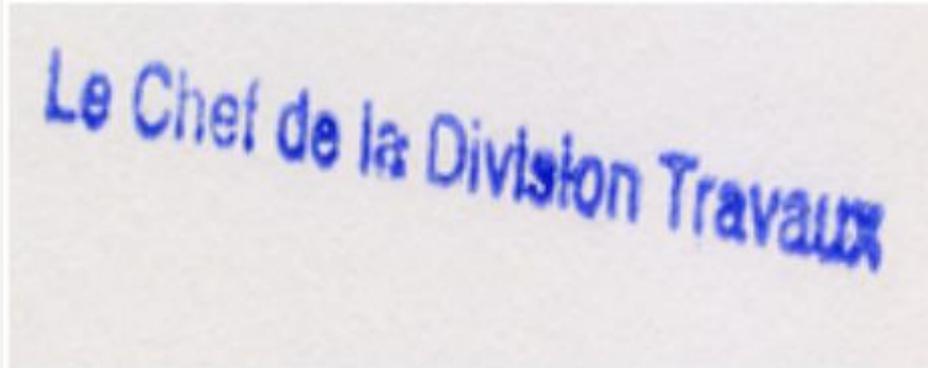
SVM CLASSIFICATION: STAMP/NON-STAMP

- Now all the candidates are segmented from the last step, those candidates contain all kinds of non-stamp candidates.

- Support Vector Machine is used for both classification models in each of two stages presented above. SVM is used, the regularizer is applied in the final output layer and the hinge loss is used as a loss function.



SVM CLASSIFICATION: GENUINE STAMP/FORGED STAMP



Genuine stamp



Forged stamp

- This thesis uses several feature extractors from the scikit image library for feature extraction. The purpose is to test various scikit image feature extractor and select the ones that produce the best result.
- GridSearchCV is applied to calculate the optimum parameter C and gamma. The best parameters are {'C': 10.0, 'gamma': 1e-09} .
- For stamps, there are many different features. In many cases, using a certain features extractor doesn't help to improve. Therefore, it is necessary to determine which features extractors to combine to achieve the best result. Various classification models is applied and implemented.

EXPERIMENT AND RESULT

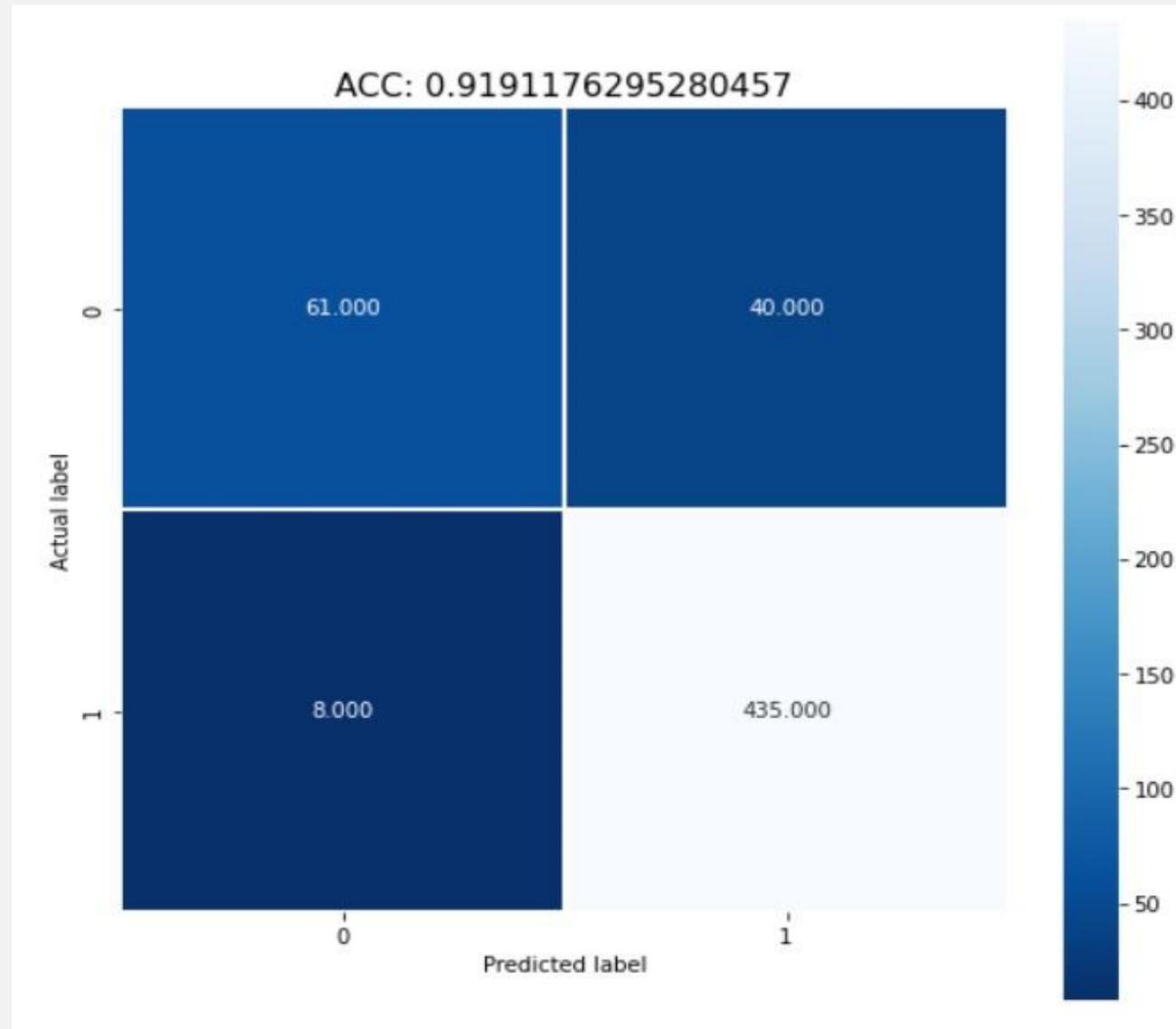
EVALUATION

- This work evaluated both classification stages, detection and verification, separately with the public data set StaVer which contains 400 document pages for evaluation of stamp detection algorithms.
- To evaluate the performance of verification, the data set is split into training set and test set with the ratio 8:2. All of the models in this work are trained and tested on these two sets.

RESULT



Performance of first SVM model on training set



Confusion matrix of first SVM model on test set

RESULT

Classifier name	Accuracy
Logistic Regression	0.93
SVM, adj	0.82
SVM, linear	0.94
SVC	0.96
K-nearest neighbors	0.89
Decision Tree	0.91
Random Forest	0.97
Random Forest 2	0.82
AdaBoost	0.93
Naive Bayes	0.84
Gradient Boosting	0.93
Latent Dirichlet allocation	0.94

Table: Result by classifiers

CONCLUSION

- This thesis introduces a simple and efficient approach to the stamp verification problem. We choose the process of 3 stages: Segmentation, classification stamp or non-stamp, classification forged stamp or genuine stamp.
- In the first two stages, we apply basic and easy to implement methods and models. The final stage is where we apply, compare various methods and algorithms to come up with the best combination of features extractors and classification models.



FUTURE WORK

Future work: Develop this approach in more kinds of stamps, the idea is to shift the focus to Vietnamese stamps.





PATIENT PRIVACY AND CONSENT (Read and sign below):

The information you provide will be used by _____ or Assistance Foundation and parties acting on their behalf to determine eligibility, to manage and improve the _____ program, products and services, to communicate with you about your experience with the _____ program, and/or to send you materials and other helpful information and updates related to the program.

By signing below, I affirm that my answers and my proof-of-income documents are complete, true and accurate to the best of my knowledge. I understand that:

- Completing this enrollment form does not guarantee that I will qualify for _____
- _____ accuracy of the information I have provided and may ask for more financial and insurance information.
- Any medicines supplied by the _____ program shall not be sold, traded, bartered or exchanged.
- _____ reserves the right to change or cancel the _____ program, or terminate my enrollment at any time.
- The support provided in this program is not contingent on any future purchase.

I certify and attest that if I receive medicine(s) provided by _____ program(s):

- I will promptly contact _____ if my financial status or insurance coverage changes.
- I will not seek to have this medicine or any cost from it counted in my Medicare Part D out-of-pocket expenses for prescription drugs.
- I will not seek reimbursement or credit for the medicine(s) from my prescription insurance provider or payor, including Medicare Part D plans for any costs of medications.
- I will notify my insurance provider of the receipt of any medicines through _____
- I have a signed copy of a current and completed HIPAA Authorization Form on record with my Prescriber so that my Prescriber may share health information about me with the _____.

Signature of Patient (Parent or guardian, if under 18 years of age) *x Virginia R. Smith* Date: *11/28/2016*

RL *131*

2016570667

DEC 19 2016

Central Standard Time | SVR:BLNMS01/44 | DNS:10107 | CSID: | DURATION (mm-ss):02:50

The data set in this work relies entirely on the published data set StarVer. We would like to construct and publish our own stamp verification data set containing only Vietnamese stamps.

Thank you for
listening