On the feasibility of classification-based product package authentication

R. Schraml, L. Debiasi, C. Kauba, A. Uhl

Department of Computer Sciences

UNIVERSITÄT SALZBURG

content overview

Motivation [1]
Counterfeited products

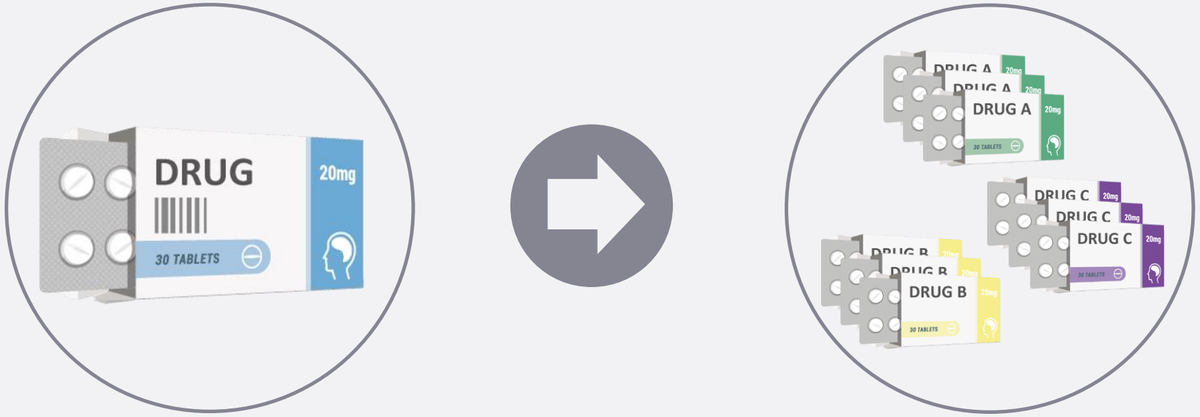
2013: *5% counterfeited products on EU level → faked medicals are a threat for the patients and cause an economic loss.*

The Falsified Medicines Directive (FMD) should be implemented until 2018. The approached solution is relies on product ***serialization*** *and tracking using unique numeric identifiers.*

Motivation [2]
Paper-based PUFs

Previous literature showed that the fibre structure of paper or packaging material is positional highly unique and enables to identify single instances.

Basic idea
move from serialization to classification



serialization

Individualize each instance of a product using unique identifiers or PUF-based approaches, e.g. fibre fingerprints

classification

Use intrinsic or extrinsic features which are constant across all instances but different to features from other products.

Pre-requirements
Fundamental research questions

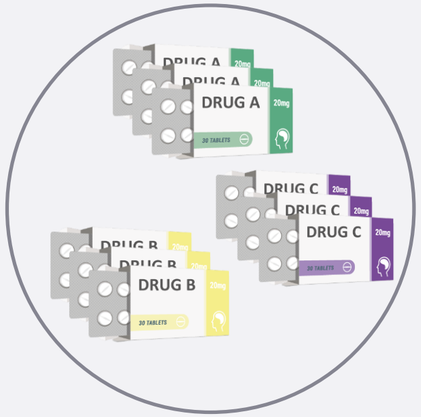
It is clear that the fibre-structure is locally unique.

Uniqueness

serialization

Individualize each instance of a product using unique identifiers or PUF-based approaches, e.g. fibre fingerprints

Pre-requirements
Fundamental research questions



classification

Use intrinsic or extrinsic features which are constant across all instances but different to features from other products.

01

It is not clear
- If the fibre structure shows constant features across different regions and
- if those features are discriminative enough to distinguish between different types of paper or packagings.

Instance invariance

?

02

It is not clear
- how the texture and the computed features vary between different instances of a product.

Instance generalisation

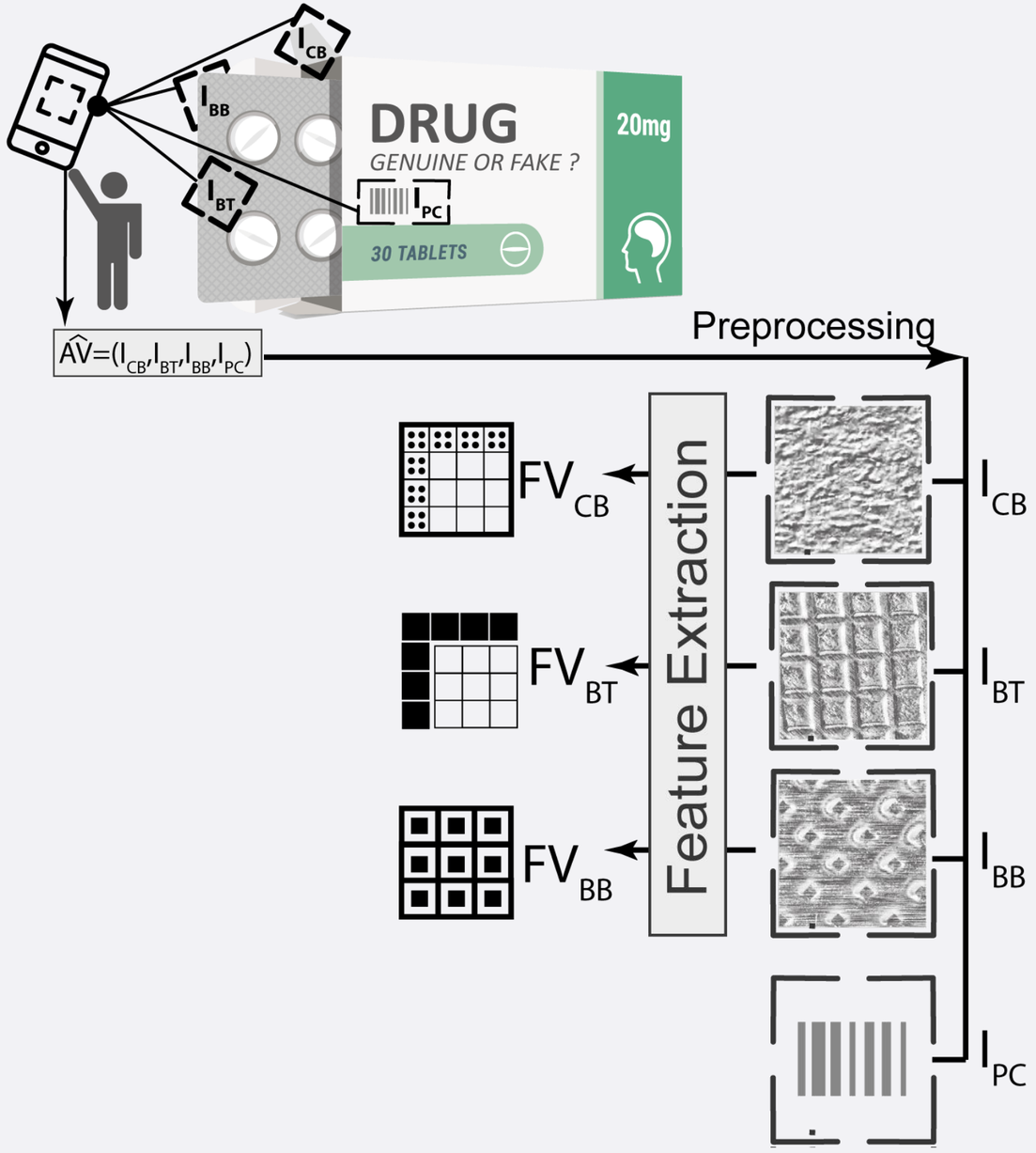
Drug packaging authentication system
Basic concept [1/3]

Capture packaging modalities:

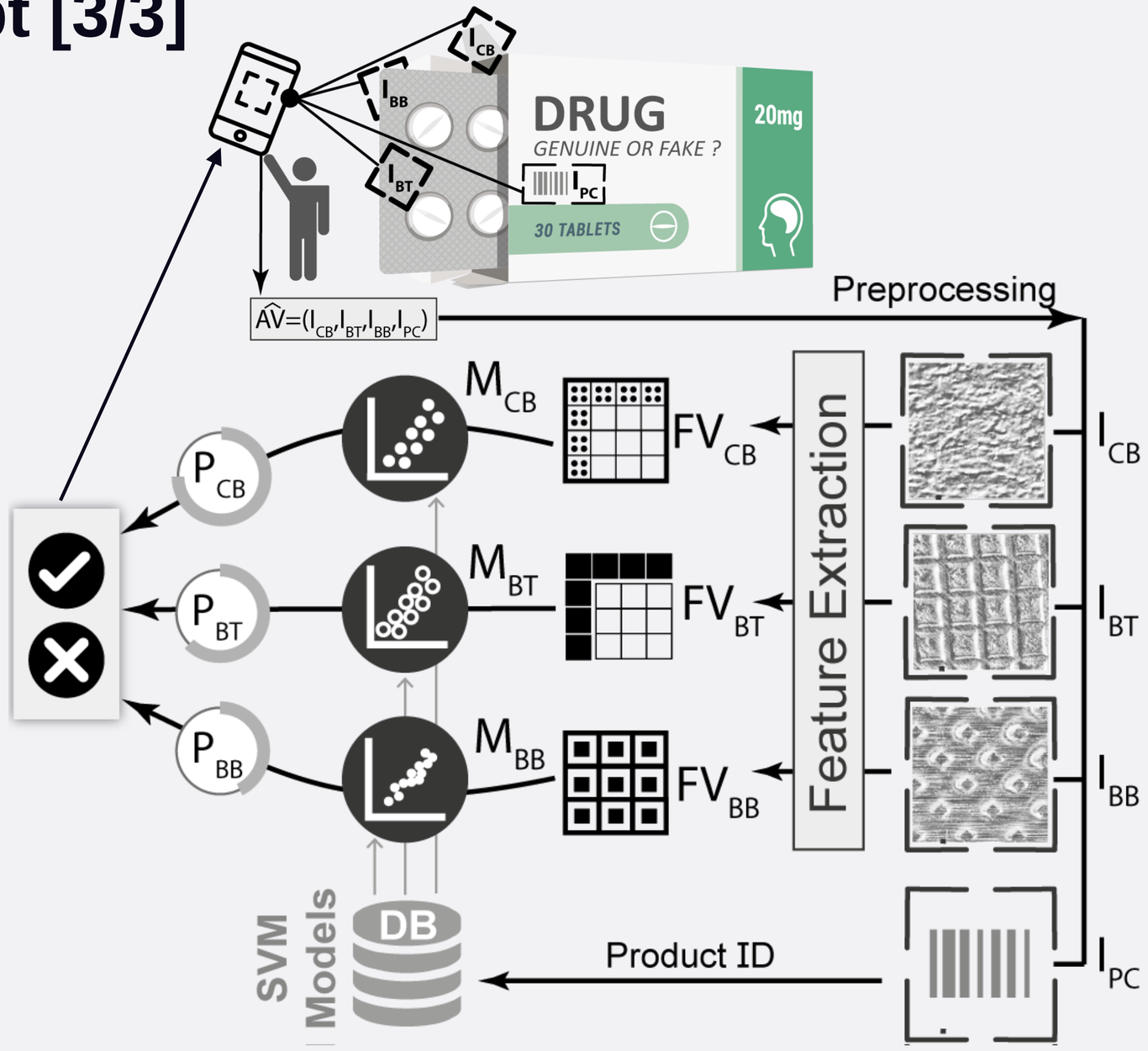
CB = Cardboard
BB = Blister Bottom
BT = Blister Top

& the product code (PC)

Drug packaging authentication system
Basic concept [2/3]



Drug packaging authentication system
Basic concept [3/3]



Drug packagings texture database
Acquisition details [1/2]

Sample collection

Packages were collected in different pharmacies in Salzburg.

Sorting & Labelling

All packages were sorted and each drug was assigned an identifier and the available instances were numbered.

- **Drugs #45**
- **Producers #28**
- **1 to 15 instances per drug**

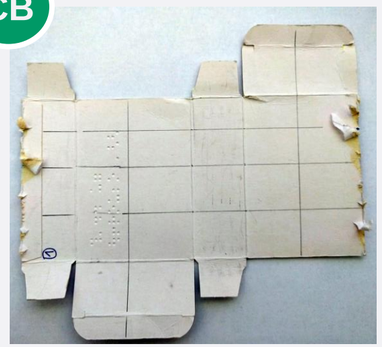
Image Acquisition

Images were captured in a controlled environment using a Canon 70D with a 100 mm macro lens and a flashlight.

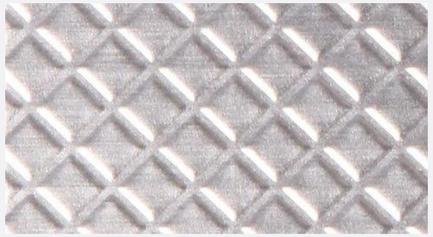
The image distance was approximately 28cm.

Drug packagings texture database
Acquisition details [2/2]

CB



BT



BB

1

Non-overlapping

Capture non-overlapping sections of each instance and modality

2

Cropping

The final images are of arbitrary size and show textural information of the modality.

Classification pipeline
Scenarios

Scenario

CLASS or PACKAGE

01

It is not clear
- If the fibre structure shows constant features across different regions and
- if those features are discriminative enough to distinguish between different types of paper or packagings.

Instance invariance

CLASS

02

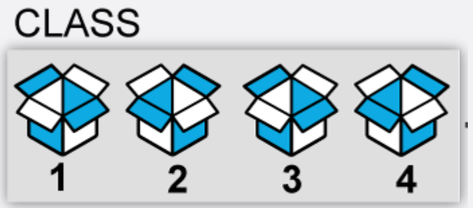
It is not clear
- how the texture and the computed features vary between different instances of a product.

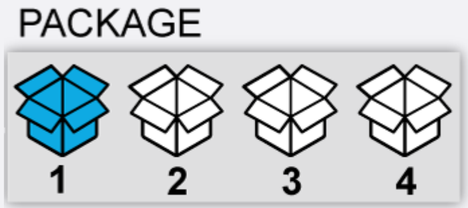
Instance generalisation

PACKAGE

Classification pipeline
Data selection

Scenario → **Data**

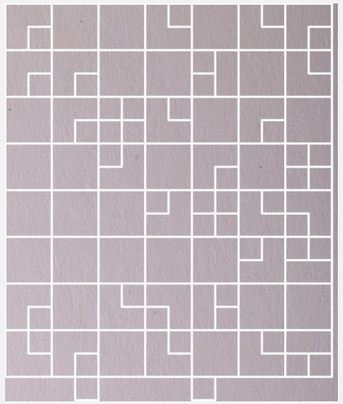


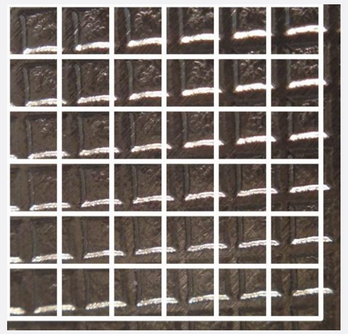


CLASS or PACKAGE

Keypoint selection:
k – image patches, with a predefined size, are selected for each modality in a scenario specific manner.

Patch sizes:
128x128, 256x256





Classification pipeline
Feature Extraction

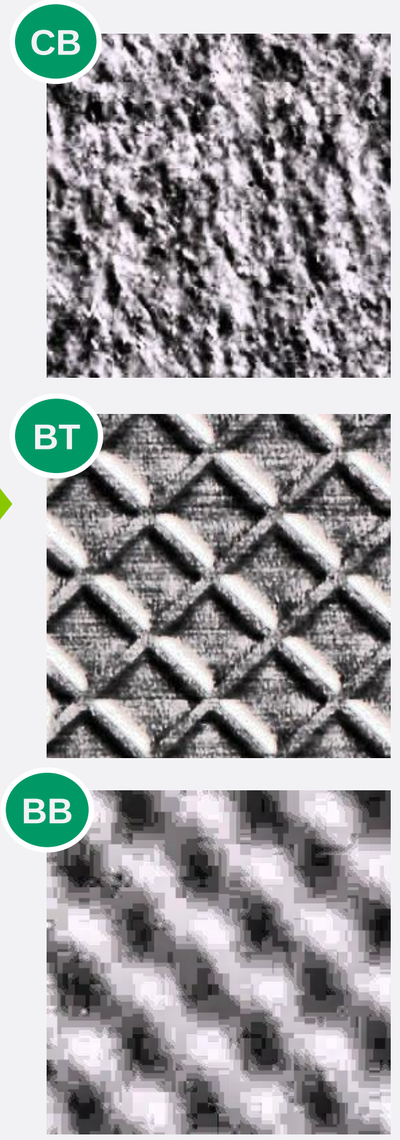
Scenario → **Data** → **Feature Extraction**

CLASS or PACKAGE

Keypoint selection

Low-level features:
Each selected patch is contrast enhanced (CLAHE) and a set of feature vectors are computed.

e.g. LBP & variants, HOG, DTCWT



Classification pipeline
Classification approaches

CLASS or PACKAGE

Keypoint selection

Low-level features

Applied SVMs:

- LIBSVM:
  - **ONE-CLASS**
  - **BINARY C-SVC**
- LIBSVM Extension:
  - **WSVM**

For all a radial basis function (RBF) kernel is utilized.

Classification pipeline
Cross-fold validation [1/4]

CLASS or PACKAGE

Keypoint selection

Low-level features

Applied SVMs

Data partitioning, hyperparameter optimization and evaluation

Classification pipeline
Cross-fold validation [2/4]

Scenario → Data → Feature Extraction → Classification Approach → Cross Validation

Parameters	
Drugs #45	D = {d_1,…,d_{45}}
Drug manufacturers	DM = {dm_1,…,dm_{28}}
Packaging modality	M = {CB, BB, BT}
Feature Extraction M.	FE = {fe_1,…, fe_n}
Classification Scenario	CS = {CLASS,PACKAGE}

$$CC = (d \in D, m \in M, fe \in M, cs \in CS\}$$

Classification Configuration

$$FV_{CC} = \{ FV_{(d_1,m,fe,cs)},..., FV_{(d_{45},m,fe,cs)} \}$$

CC specific Feature Vector Sets

Nested cross-validation using a specific classification approach

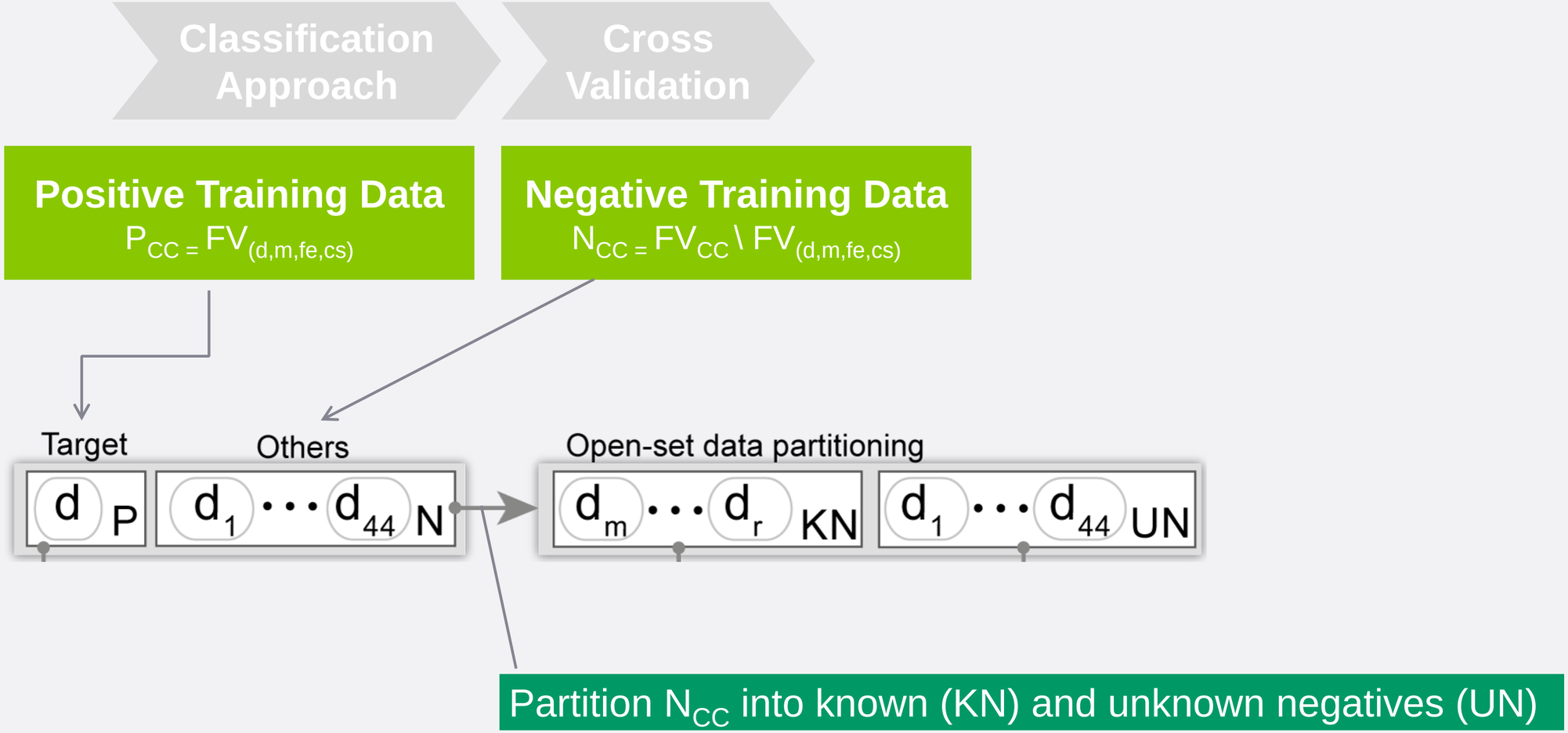
Positive Training Data

$P_{CC} = FV_{(d,m,fe,cs)}$

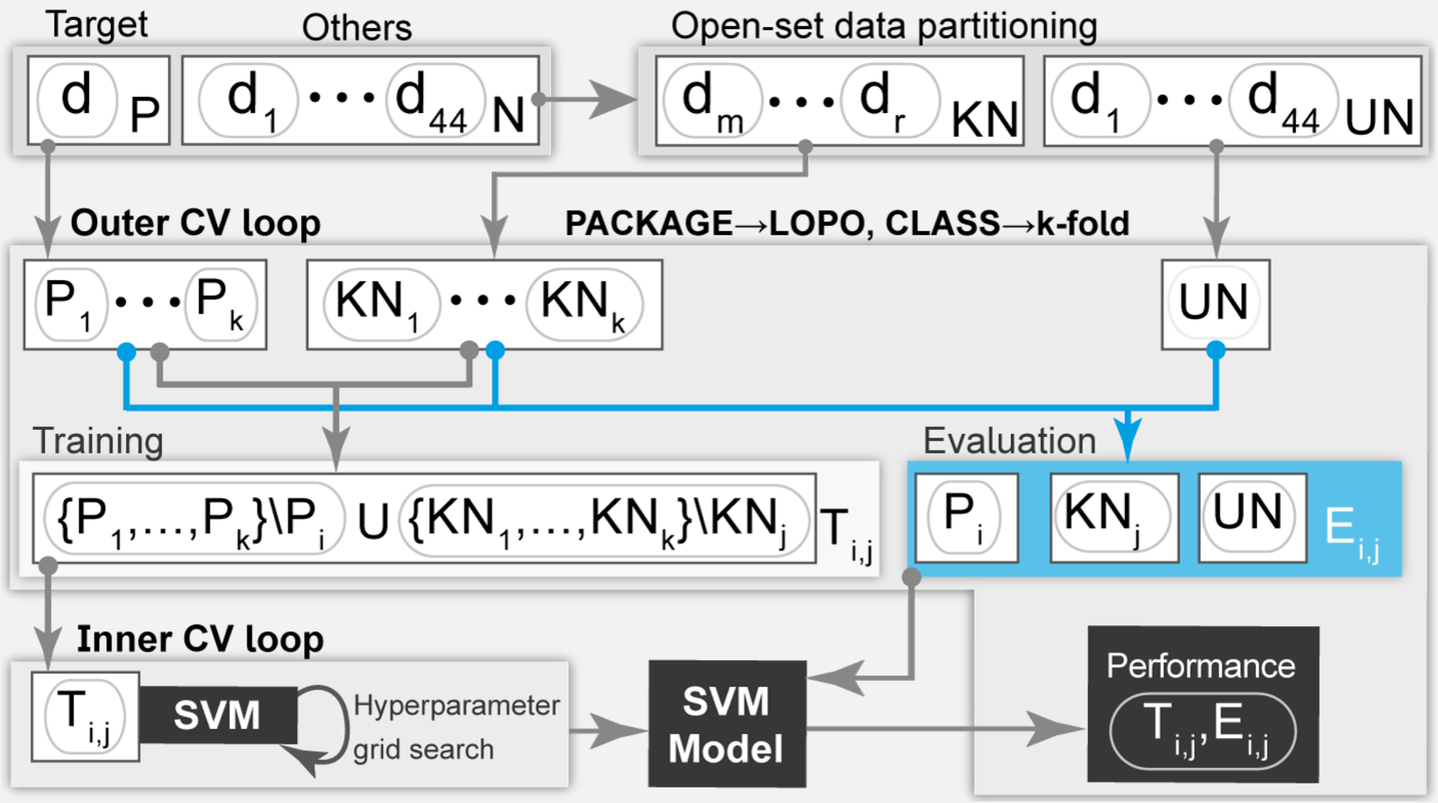
Negative Training Data

$N_{CC} = FV_{CC} \setminus FV_{(d,m,fe,cs)}$

Classification pipeline
Cross-fold validation [3/4]



Classification pipeline
Cross-fold validation [4/4]



Optimize classifier using known negatives in the inner CV loop

In case of binary SVMs, a subset of the known negatives is not used for training; i.e. only for evaluatuin in order to address the open-set problem in the inner CV loop. The SVM parameters and a probabilitay threshold achieving the highest F-Measure are determined.

Experiments
Research questions - Reminder

01

It is not clear
- If the fibre structure shows constant features across different regions and
- if those features are discriminative enough to distinguish between different types of paper or packagings.

Instance invariance

CLASS

02

It is not clear
- how the texture and the computed features vary between different instances of a product.

Instance generalisation

PACKAGE

Classification Performances – Overview:

All CCs were computed and the best CLASS and PACKAGE configurations (=highest averaged F-Measure and StDev) for different CC parameters and modalities were determined.

- **CLASS:** averaged results over all 45 drugs
- **PACKAGE:** averaged results for drugs with at least 5 instances.

Experiments
Results overview

01

02

page 24

B

	CLASS						PACKAGE					
	128×128			256×256			128×128			256×256		
CC	CB	BT	BB	CB	BT	BB	CB	BT	BB	CB	BT	BB
ONE-CLASS	*LTP* 0.83 ±7.9	*LTP* 0.9 ±6.2	*LTP* 0.92 ±5.8	*LTP* 0.91 ±4.4	*LTP* 0.85 ±13.6	*LBP* 0.87 ±13.5	*LBP* 0.81 ±8.7	*LBP* 0.86 ±6.3	*LTP* 0.84 ±11.3	*LTP* 0.85 ±9.1	*LBP* 0.88 ±5.0	*LBP* 0.85 ±7.1
BINARY	*LTP* 0.88 ±6.9	*LiLBP* 0.94 ±3.2	*LTP* 0.93 ±4.1	*LTP* 0.91 ±5.2	*LiLBP* 0.92 ±9.0	*LTP* 0.93 ±5.0	*LTP* 0.82 ±9.5	*LTP* 0.92 ±3.7	*LTP* 0.87 ±8.9	*LTP* 0.85 ±5.5	*LTP* 0.94 ±5.7	*LiLBP* 0.87 ±10.0
WSVM	*LTP* 0.86 ±7.6	*LTP* 0.93 ±4.1	*LTP* 0.93 ±4.3	*LiLBP* 0.88 ±6.0	*LTP* 0.88 ±7.6	*MFS* 0.88 ±9.1	*LTP* 0.85 ±8.2	*LTP* 0.91 ±4.2	*LiLBP* 0.85 ±9.2	*LiLBP* 0.83 ±8.5	*LTP* 0.89 ±8.7	*LiLBP* 0.84 ±10.1

page 23

A

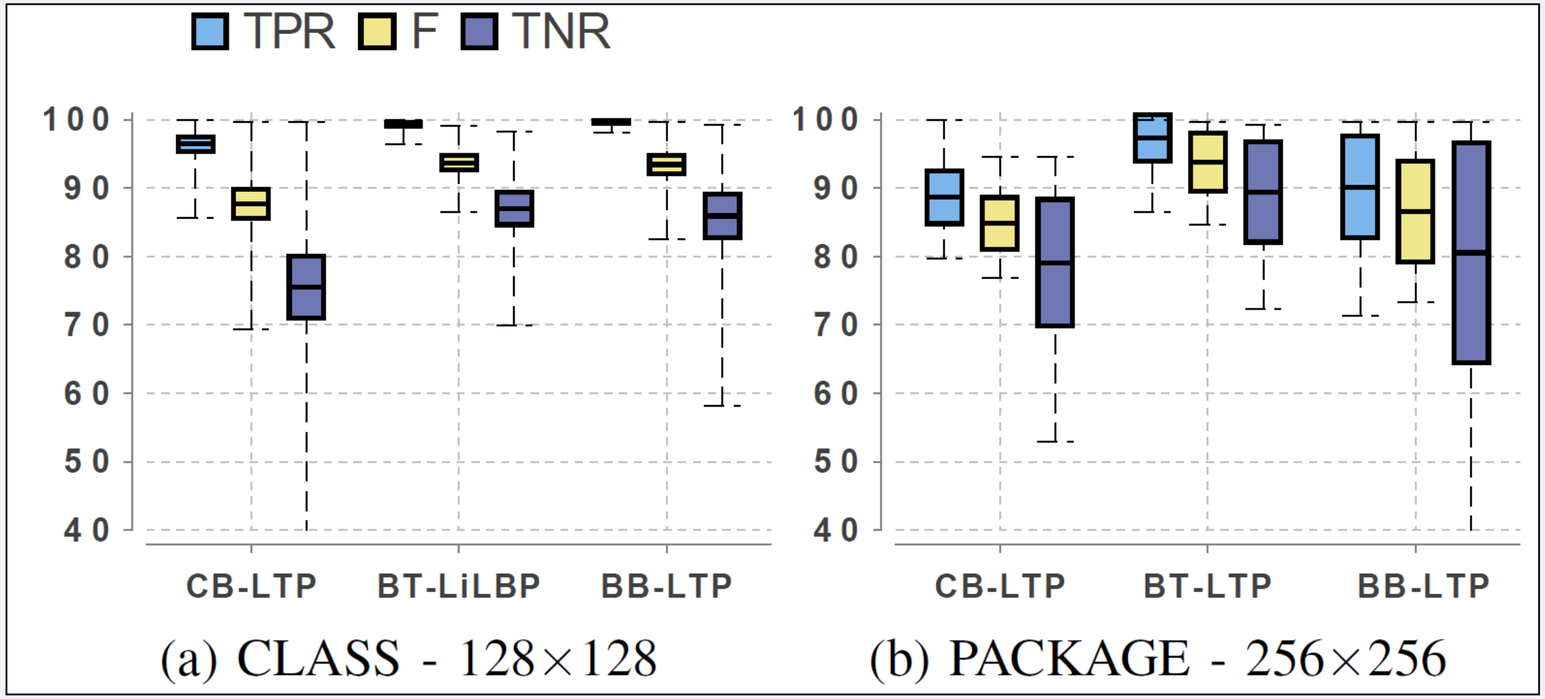
CLASS results show high mean F-Measures over 0.9, indicating that textures from all three modalities show constant but highly discriminative features. This enables to recognize the same drug class and to distinguish it from others.

Instance invariance

Package results show that textural features are constant across different instances for all three modalities.
This is a basic requirement for a
classification-based authentication system.

Instance generalisation

Experiments
Details [A]



Performance (StDev) decreases in case of the PACKAGE scenario which reflects a real world package-authentication setup.

CLASS vs. PACKAGE (Binary C-SVC) *performance comparison.*

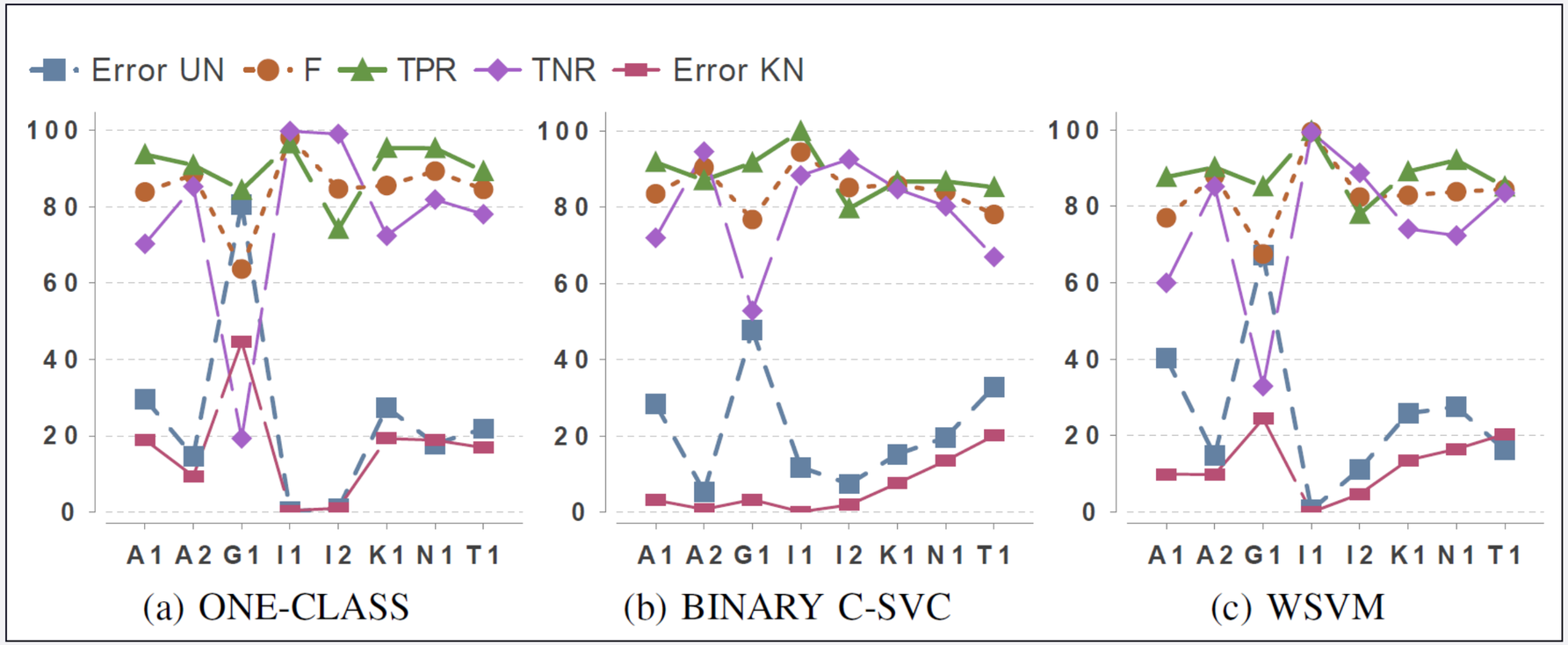
Class Accuracy / True Positive Rate:

$$TPR = \frac{TP}{TP + FN}$$

Others Accuracy / True Negative Rate

$$\mathrm{TNR} = \frac{TN}{TP + FP}$$

Experiments
Details [B]



PACKAGE (256x256) - SVM performance comparison *for **CB** and all target drugs (=8) with more than 5 instances.*

X-Axis: Target Drug IDs: (e.g. A1 = manufacturer A + drug number 1.

Error UN = False classified unknown negatives.

Error KN = False classified known negatives.

Error KN (data seen in training) is lower than Error UN (=data not used for training = open world).

Thank you
Conclusions and Outlook

Instance invariance

Textural features of drug packaging material are constant and highly discriminative.

Instance generalistation

Experiments indicate that a classifier can be trained with a set of known instances and is able to authenticate unseen instances.

TODO

Real world data from faked packages is required. Use high-level features, feature encoding and feature fusion techniques. Investigate error sources, i.e. probably other drugs from the same manufacturer use the same packaging material.