

Expert System for Insurance Data Fraud Detection

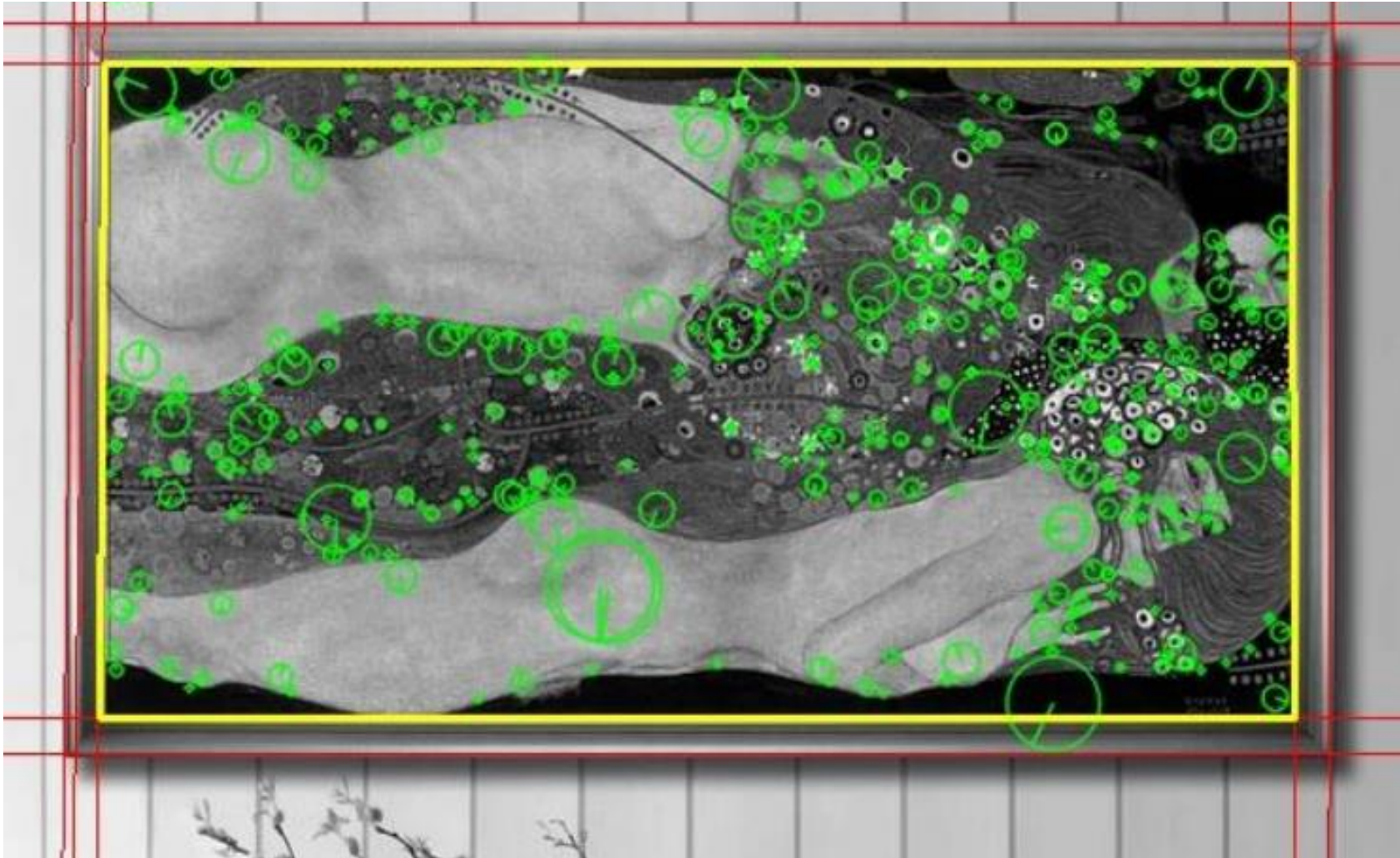
Jaroslav Ráček, Josef Daňa

TECHNISERV, spol. s r.o.
Masaryk University, Faculty of Informatics

FRAUDES System

- Tool for insurance frauds detection
- Combines 2 types of analyses
 - Image analysis – property and vehicle insurance
 - Text analysis – life insurance
- Supportive functionality based on neuronal networks

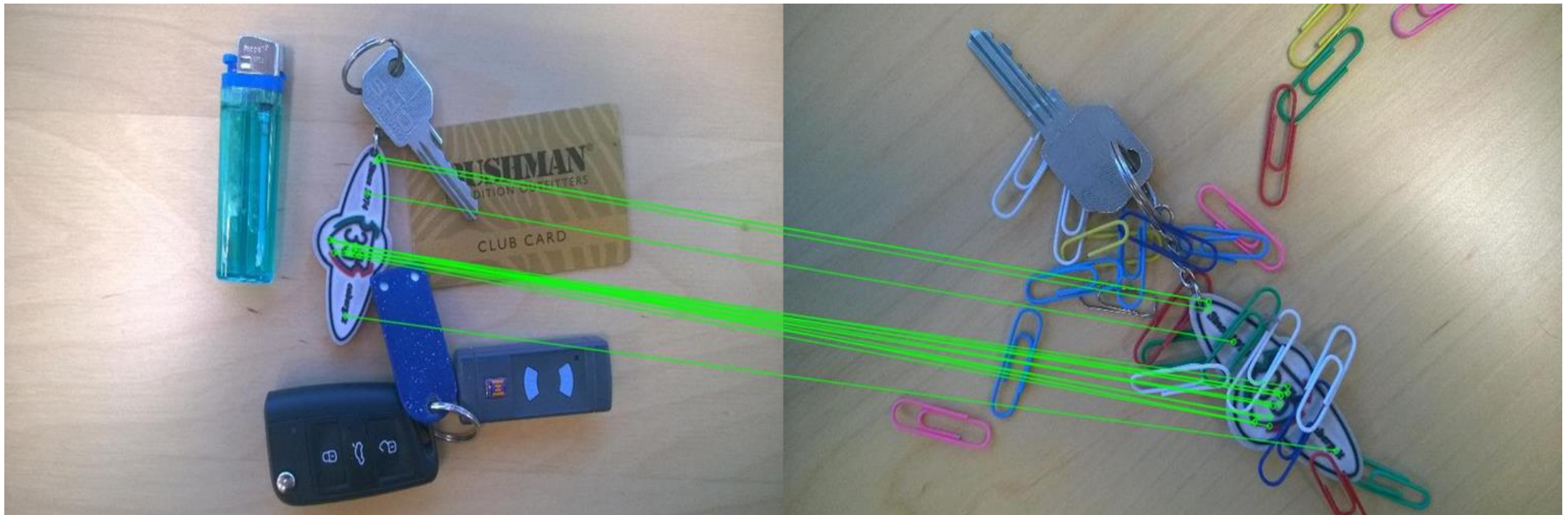
Local Descriptors



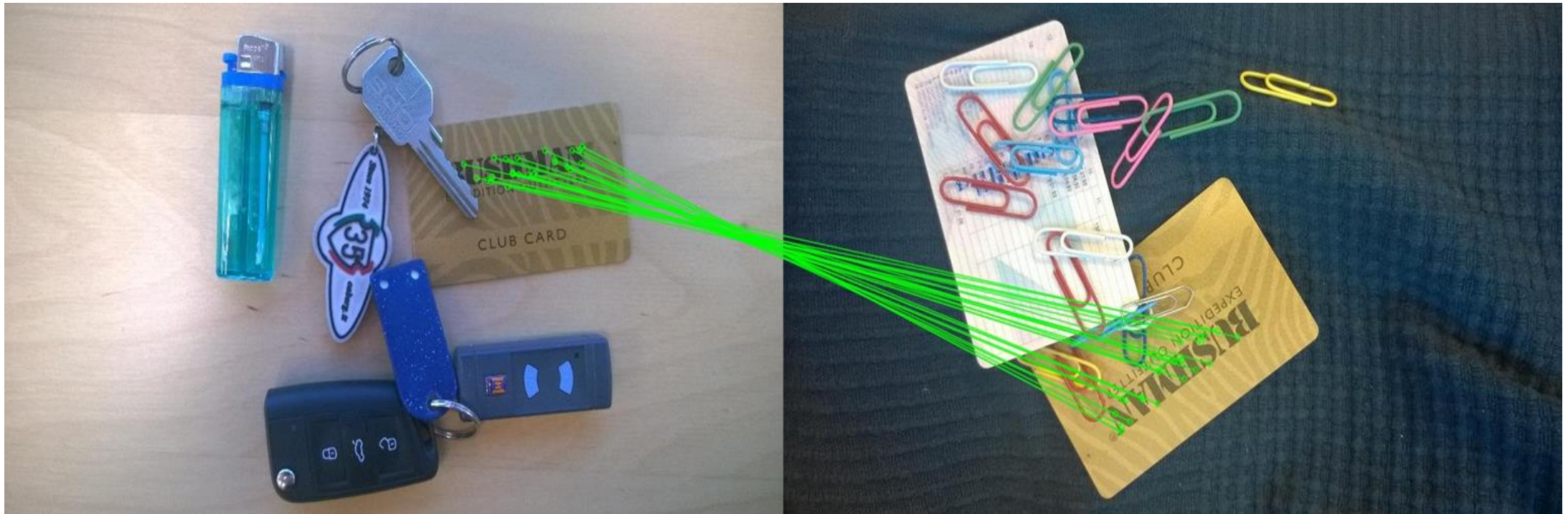
Used Descriptors

- HOG (Histogram of Oriented Gradient) – robust but slow
- Binary descriptors – fast but not very robust
- SIFT – used in first version of the system
- KAZE and AKAZE – current development

Identification Process of Our Solution



Identification Process of Our Solution



Properties of Our Tool

- Nearby objects does not matter
 - Background and nearby (irrelevant) objects are successfully separated and the analysis discover identical subject
- Discovering partially covered object
 - Keypoints are mutually independent. There is a chance that partially covered object will be succesfully identified based on the visible keypoints
- Rotation and different scale does not matter
 - Keypoints keep their characteristics even after rotation or zoom. Consequently, an object can be identified from different angles or distances

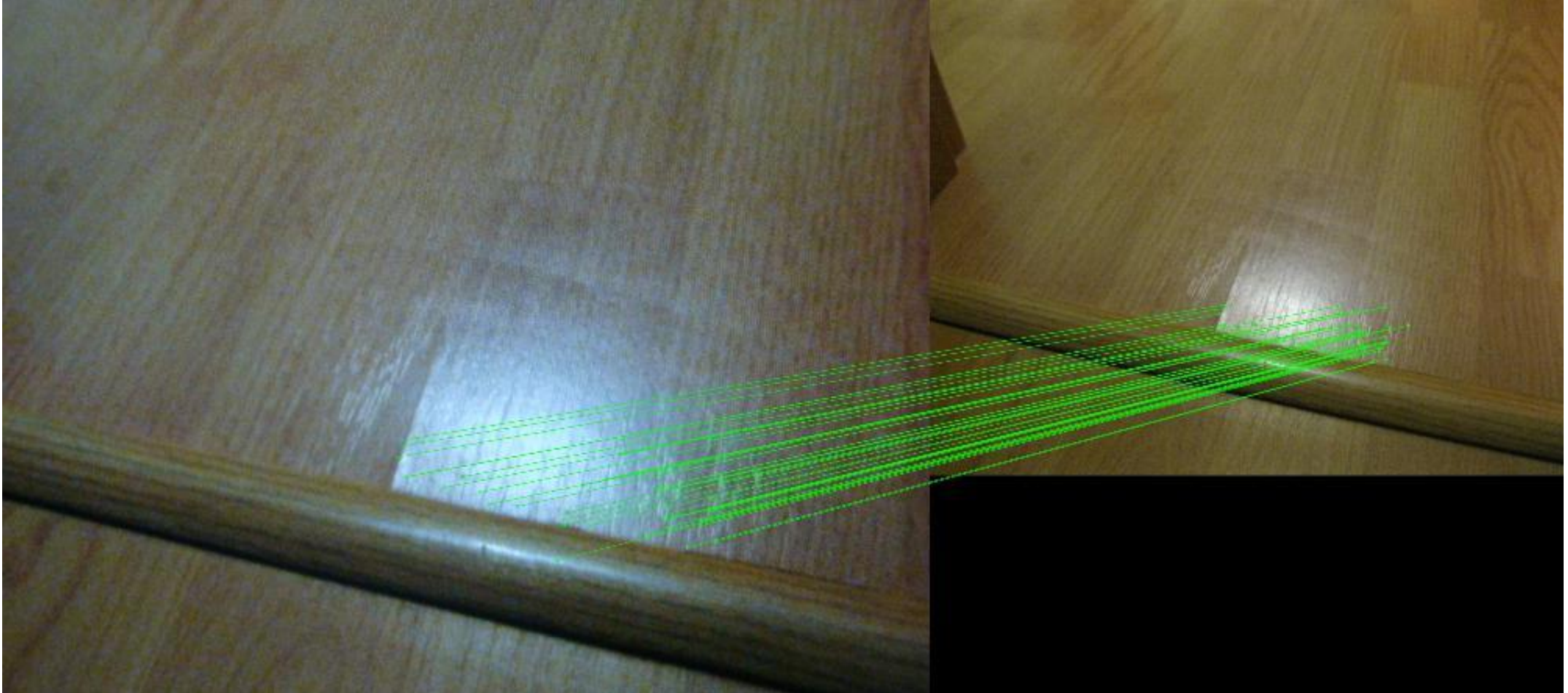
Examples of Insurance Company Data Analysis (1)



Examples of Insurance Company Data Analysis (2)



Examples of Insurance Company Data Analysis (3)



Robustness against Color Change (1)



Robustness against Color Change (2)



HW Requirements of Image Analysis

- At least two workstations are recommendable
 - 1 workstation for image markers computation
 - 1 workstation for image comparison/search tasks
- On a common workstation, we are able to compare approx. 25 million pairs of images per day, i.e. approx. 1 million pairs per hour
- The solution is scalable

Text analysis

- Detection of identical text segments
 - Medical records, insurance claim records
- Authorship analysis
 - Vocabulary comparison
 - “Anomaly “ detection
 - Medical records

Usage of Neuronal Networks

- Detection of damaged areas
 - Better focusing of image markers
- Identification of personal data in an image
 - Cleaning of data before the analysis

Current Appearance of the Application (1)

Home / Browse duplicate cases ✉

Show: All Advanced filters Search... Q

	Duplicity To inspect Disagreement (technical) Disagreement (factual) Inspect ▶	
? 15.20006.0017420	$\rightleftharpoons$	? 15.20006.0012074
1691.75		
	Duplicity To inspect Disagreement (technical) Disagreement (factual) Inspect ▶	
? 15.20006.0002655	$\rightleftharpoons$	? 15.20006.0002619
689.52		
	Duplicity To inspect Disagreement (technical) Disagreement (factual) Inspect ▶	
? 15.20006.0014537	$\rightleftharpoons$	? 15.20006.0011727
406.35		

Current Appearance of the Application (2)

Comparison of case [15.20006.0014537](#) with case [15.20006.0011727](#)

Show: All Advanced filters Q



6278233.jpg

[15.20006.0014537](#)



9.88



6286619.jpg

[15.20006.0011727](#)

Match rejected Match confirmed ✕ ▼





9.87



Current Appearance of the Application (3)



6278253.jpg

? 15.20006.0014537

9.82



6286639.jpg

? 15.20006.0011727

Match rejected

Match confirmed ✕



Thank You for Your Attention