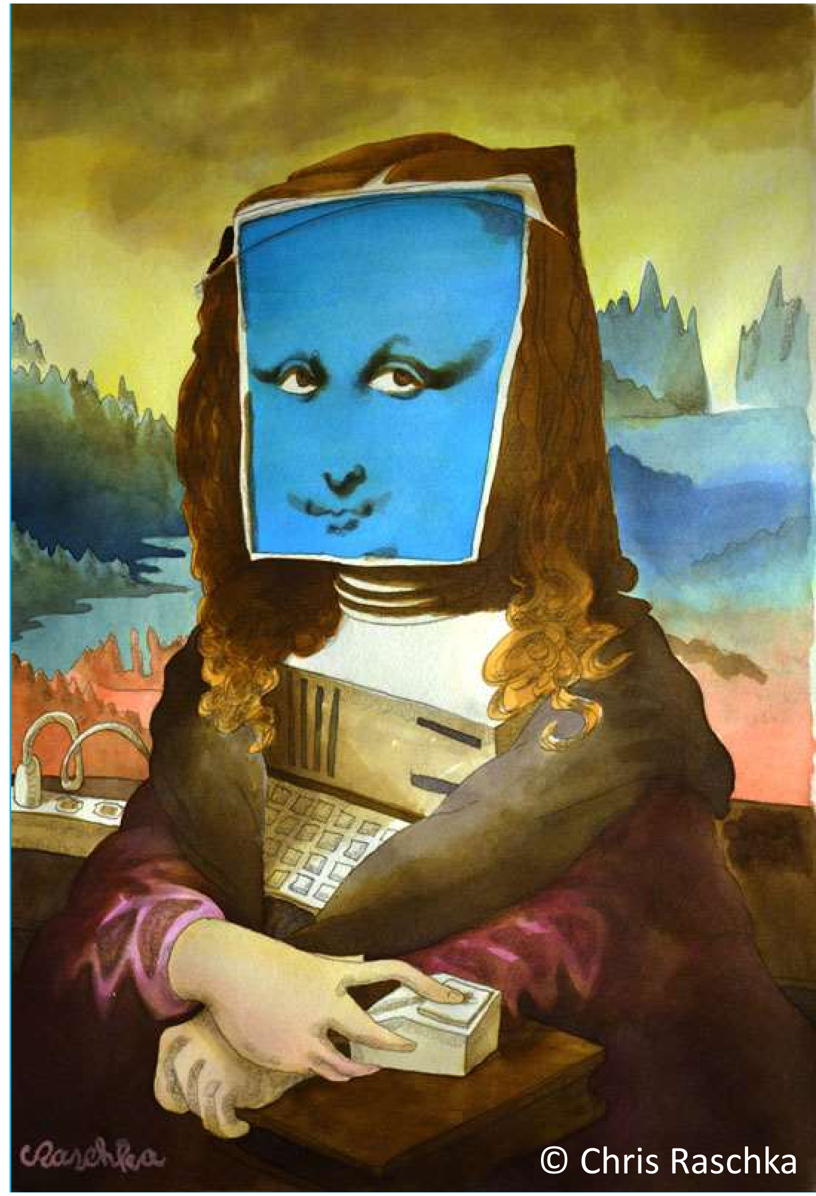
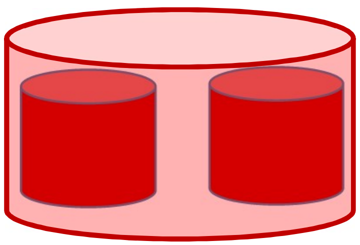
MonaLIA 1.0

Preliminary study on image recognition of the Joconde database in connection with semantic data (JocondeLab)



Study Overview

Joconde database



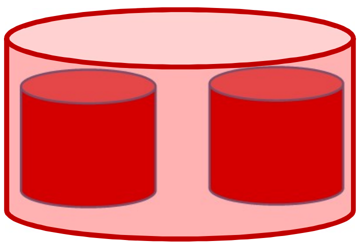
350 000 ***images*** *of artworks*

RDF ***metadata*** *based on* *external thesauri*

Study Overview

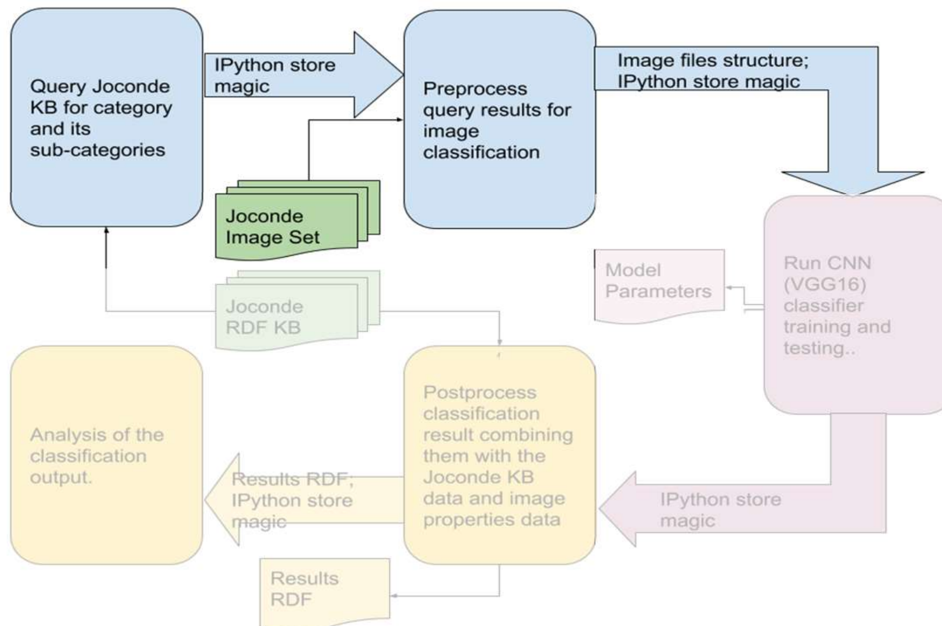
- **reason & query on RDF metadata to build balanced, unambiguous, labelled training sets.**

Joconde database



350 000 ***images*** *of artworks*

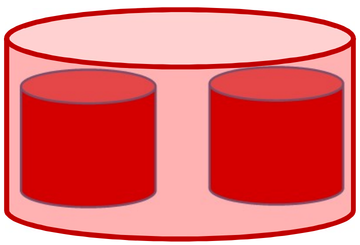
RDF ***metadata*** *based on* *external thesauri*



Study Overview

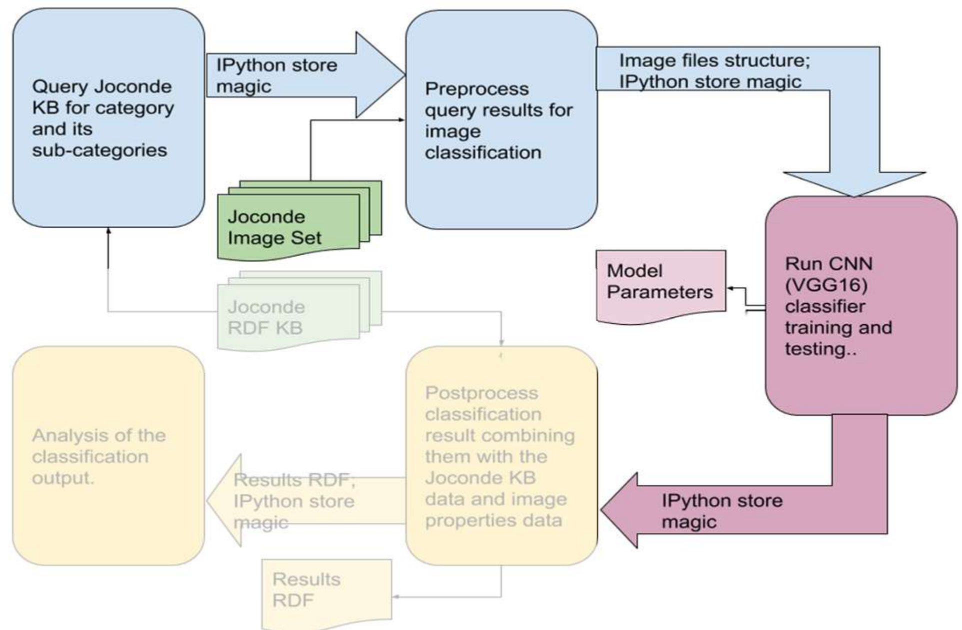
- reason & query on RDF metadata to build balanced, unambiguous, labelled training sets.
- **transfer learning & CNN classifiers on targeted categories (topics, techniques, etc.)**

Joconde database



*350 000 **images** of artworks*

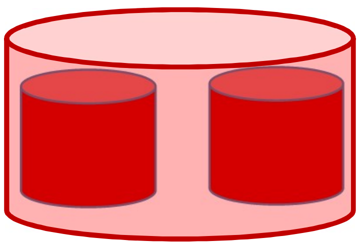
*RDF **metadata** based on <u>external thesauri</u>*



Study Overview

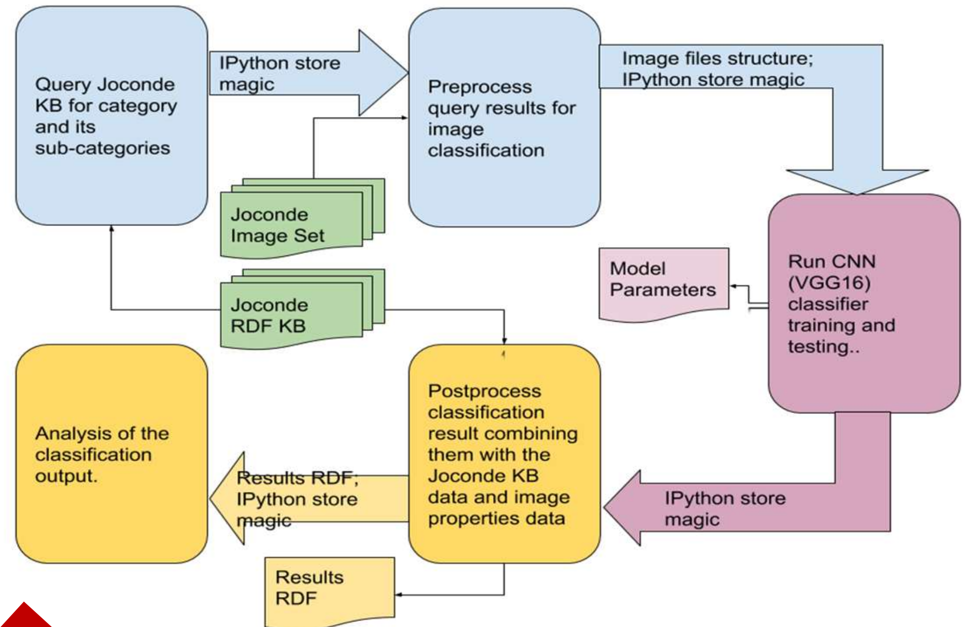
- reason & query on RDF metadata to build balanced, unambiguous, labelled training sets.
- transfer learning & CNN classifiers on targeted categories (topics, techniques, etc.)
- **reason & query RDF metadata of results to address silence, noise**

Joconde database



*350 000 **images** of artworks*

*RDF **metadata** based on external thesauri*



☑ ***animal***

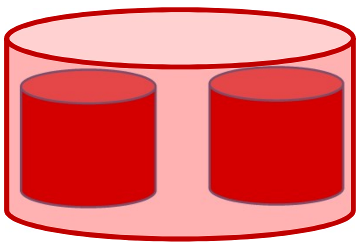
? ***bird***

☒ ***painting***

Study Overview

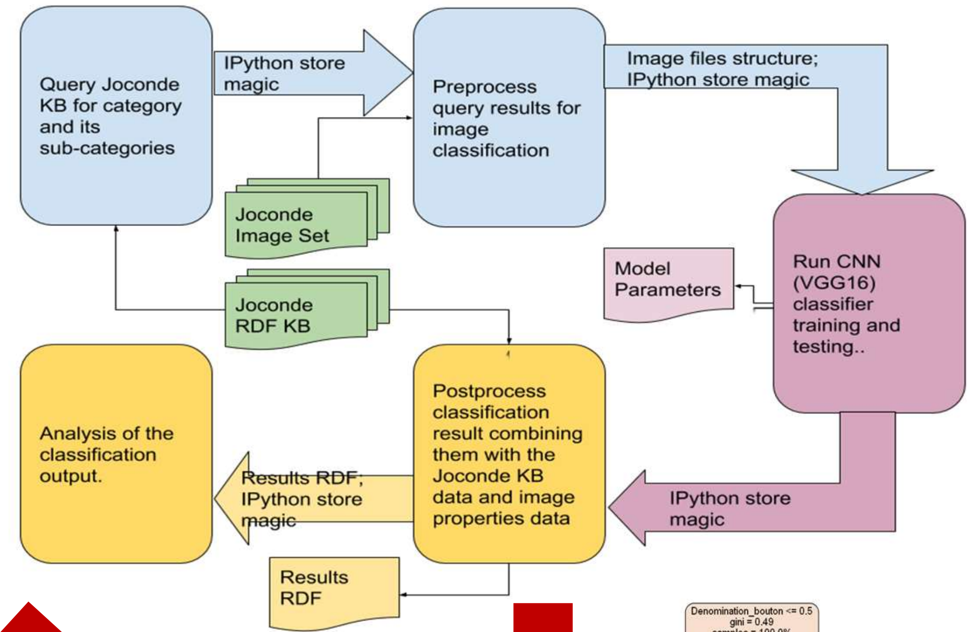
- reason & query on RDF metadata to build balanced, unambiguous, labelled training sets.
- transfer learning & CNN classifiers on targeted categories (topics, techniques, etc.)
- reason & query RDF metadata of results to address silence, noise **and explain**

Joconde database

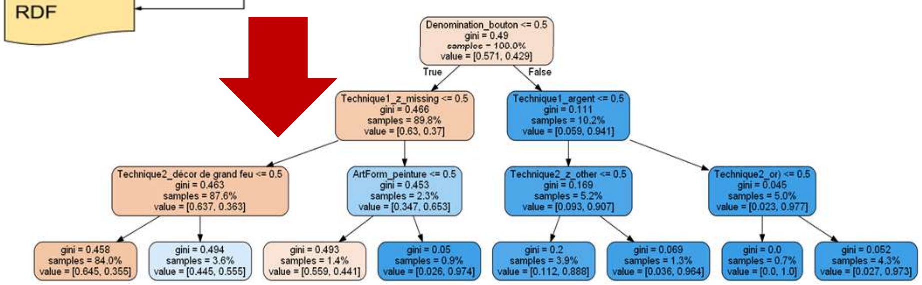


*350 000 **images** of artworks*

*RDF **metadata** based on <u>external thesauri</u>*



☑ ***animal***
? ***bird***
☒ ***painting***



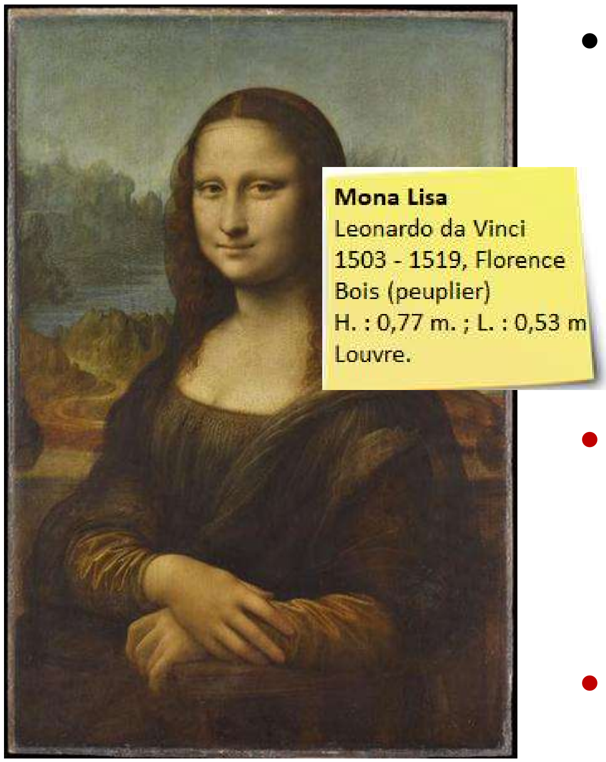
Motivation & Challenges

GALERIE DE VUES DE LA ROME MODERNE by PANNINI Giovanni Paolo

Museum curators have to annotate thousands of artworks acquired over the hundreds of years and now managed as digital collections. This process can be tedious and susceptible to the human errors and omissions.

- **Can the existing digital artwork collections be automatically enhanced by combining Machine Learning and Knowledge Representation & Reasoning?**
- **Can annotation of the new artworks be automated or semi-automated?**

Joconde Database



PORTRAIT DE MONA LISA (1479-1528) ; DITE LA JOCONDE by Leonardo Da Vinci

- 350 000 illustrated artwork records from the French museums.

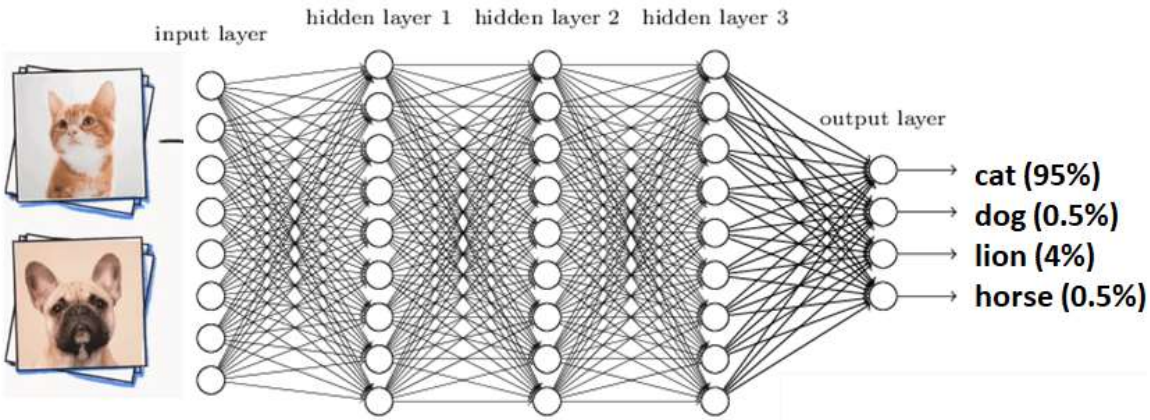
  RDF metadata describing the artwork subject and properties (media, author, museum, etc.).

- **The database is searchable on the artwork subjects and other properties but…**

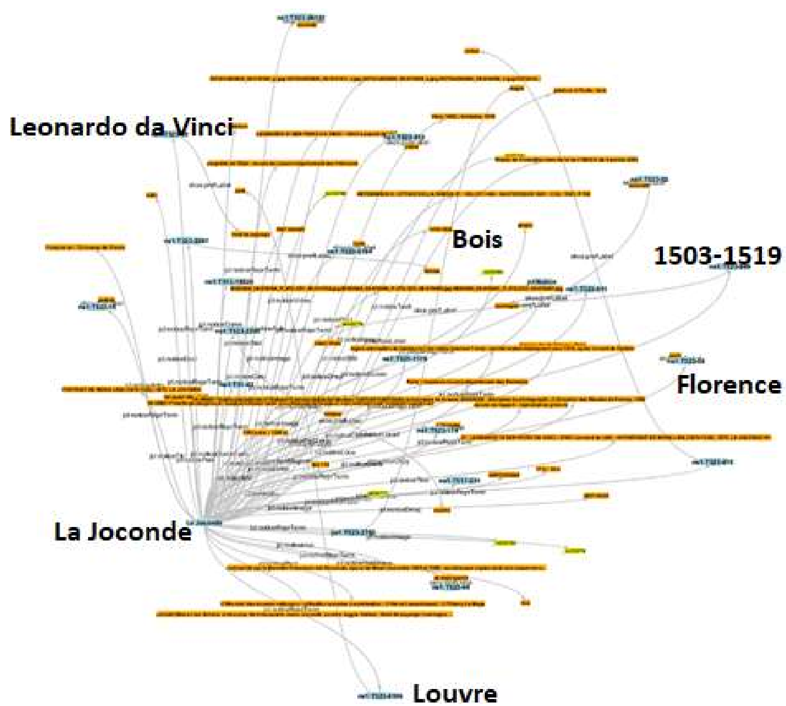
- **The metadata can be <u>incomplete</u> & <u>noisy</u>.**
- **The new artworks <u>added continuously</u>.**

Enabling Methods

Deep Learning from unstructured data such as images



Semantic Reasoning and querying from semantic metadata



Approach

- **SPARQL+RDFS+SKOS on metadata to extract training and test subset of images: query and reason to**
  - **identify class subsets with enough labeled images for training**
  - **balance number of images per class**
  - **avoid images with intersected classes (ambiguity)**

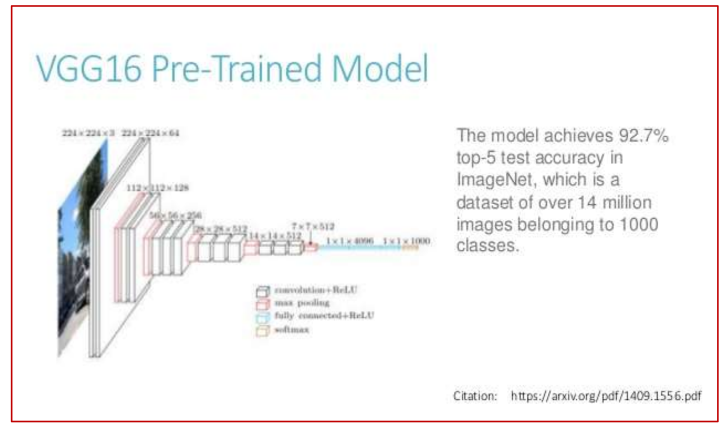
```
prefix skos: <http://www.w3.org/2004/02/skos/core#>
prefix jcl: <http://jocondelab.iri-research.org/ns/jocondelab/>

select ?categoryl ?noticeReference ?imagePath  where {
    ?category a jcl:Term.
    ?category skos:prefLabel "genre de la représentation"@fr.

    ?notice jcl:noticeReprTerm ?category.
    ?notice jcl:noticeRef ?noticeReference.
    ?notice jcl:noticeHasImage true.
    ?notice jcl:noticeImage [ jcl:noticeImageIsMain true ;
                              jcl:noticeImagePath ?imagePath].
}
```

Approach

- SPARQL+RDFS+SKOS on metadata to extract training and test subset of images: query and reason to
    - identify class subsets with enough labeled images for training
    - balance number of images per class
    - avoid images with intersected classes (ambiguity)

```
prefix skos: <http://www.w3.org/2004/02/skos/core#>
prefix jcl: <http://jocondelab.iri-research.org/ns/jocondelab/>

select ?category1 ?noticeReference ?imagePath  where {
    ?category a jcl:Term.
    ?category skos:prefLabel "genre de la représentation"@fr.

    ?notice jcl:noticeReprTerm ?category.
    ?notice jcl:noticeRef ?noticeReference.
    ?notice jcl:noticeHasImage true.
    ?notice jcl:noticeImage [ jcl:noticeImageIsMain true ;
                              jcl:noticeImagePath ?imagePath].
}
```

- **Train Deep Learning (CNN) classifier**
    - **using VGG16 to train the classifier for the artwork images**
    - **learning transfer from ImageNet to decrease training time**

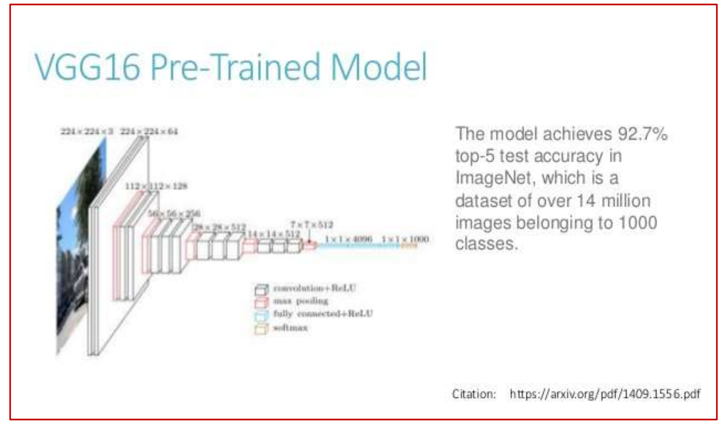


Approach

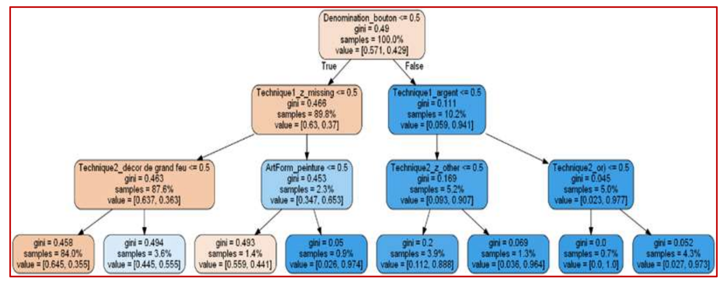
- SPARQL+RDFS+SKOS on metadata to extract training and test subset of images: query and reason to
  - identify class subsets with enough labeled images for training
  - balance number of images per class
  - avoid images with intersected classes (ambiguity)

```
prefix skos: <http://www.w3.org/2004/02/skos/core#>
prefix jcl: <http://jocondelab.iri-research.org/ns/jocondelab/>

select ?categoryl ?noticeReference ?imagePath  where {
    ?category a jcl:Term.
    ?category skos:prefLabel "genre de la représentation"@fr.

    ?notice jcl:noticeReprTerm ?category.
    ?notice jcl:noticeRef ?noticeReference.
    ?notice jcl:noticeHasImage true.
    ?notice jcl:noticeImage [ jcl:noticeImageIsMain true ;
                              jcl:noticeImagePath ?imagePath].
}
```

- Train Deep Learning (CNN) classifier
  - using VGG16 to train the classifier for the artwork images
  - learning transfer from ImageNet to decrease training time



- **SPARQL & RDF again to find dependencies between the classification outcome and the artwork properties**
  - **statistically significant variable selection**
  - **decision tree**



Results & Findings: Summary

- Although the artwork images have a great variability in quality, size and actual artwork media, the state of the art Deep Learning models can be trained to identify the depicted objects, general themes and actual artwork media.

Classifier	Number of Classes	Training Set Size	Test Set Size	Best Top-1 Test Accuracy
Animals	14	7153	893	40%
Animals vs. Humans	2	28822	2882	85%
Representation Type	4	1828	228	85%
Iconographic Type	7	6440	648	69%
Art Form	7	6377	644	83%
Art Form	5	33504	4195	94%

Results & Findings: Summary

- Although the artwork images have a great variability in quality, size and actual artwork media, the state of the art Deep Learning models can be trained to identify the depicted objects, general themes and actual artwork media.

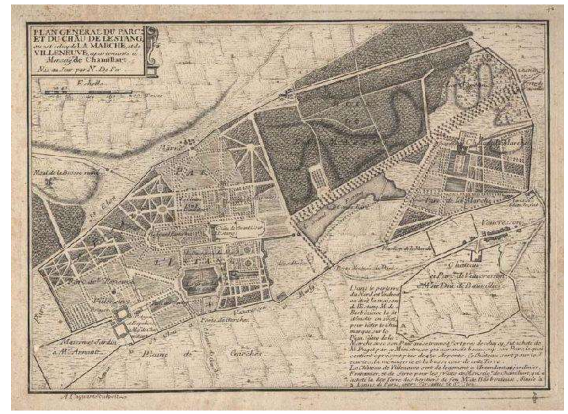
Classifier	Number of Classes	Training Set Size	Test Set Size	Best Top-1 Test Accuracy
Animals	14	7153	893	40%
Animals vs. Humans	2	28822	2882	85%
Representation Type	4	1828	228	85%
Iconographic Type	7	6440	648	69%
Art Form	7	6377	644	83%
Art Form	5	33504	4195	94%

Classifying Representation Types

- 4 classes:
    - genre iconographique
    - ornamentation
    - représentation non figurative
    - représentation scientifique
- 85% accuracy

- Correct classification
    - Label: représentation scientifique
    - Prediction: représentation scientifique with prediction score 66.98%

- Incorrect classification
    - Label: représentation scientifique with prediction score 17.93%
    - Prediction: ornamentation with prediction score 80.87%



Classifying Animals vs. Humans

- 2 classes, 85% accuracy
- Correct classification
  - Label: âge de la vie (human)
  - Prediction: âge de la vie
    with prediction score 98.96%

  - Label : espèce animale (animal)
  - Prediction: espèce animale
    with prediction score 99.25%

- Incorrect classification
  - Label : espèce animale
    with prediction score 12.83%
  - Prediction: âge de la vie
    with prediction score 87.16%

Classifying Animals vs. Humans

- Sometimes the images are mislabeled
  - Label: espèce animale (Animal)
  - Prediction: âge de la vie (Human) with prediction score 84.74%

- But misslabeling cannot always be detected
  - Label: espèce animale
  - Prediction: espèce animale with prediction score 87.93%

Classifying Animals

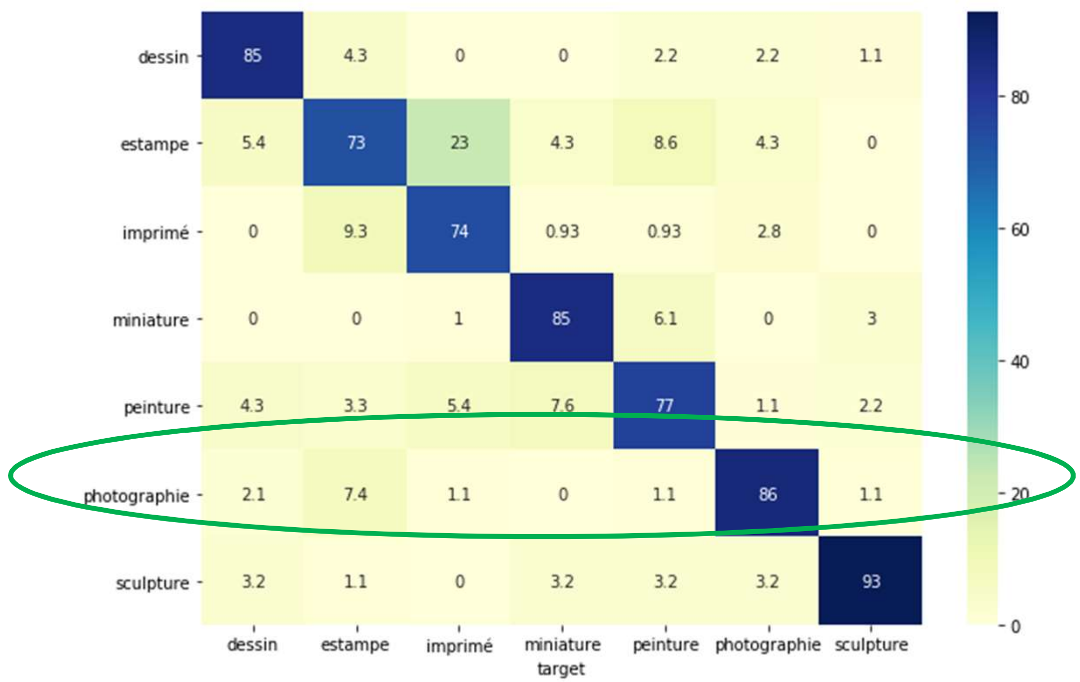
- 14 classes:
  - aigle
  - cerf
  - chat
  - cheval
  - chien
  - chèvre
  - colombe
  - lion
  - mouton
  - ophidien
  - papillon
  - sanglier
  - vache
  - âne
- 40% accuracy
- Uncertain classification (most often the case)
  - Label: chèvre
  - Prediction:

vache	27.64%
cheval	22.46%
chèvre	14.82%
mouton	13.08%
âne	10.25%

Classifying Art Forms

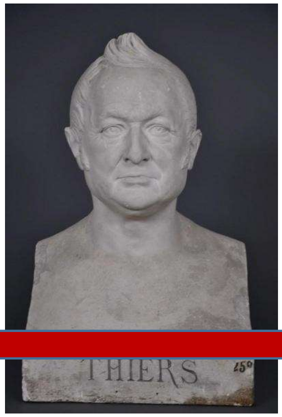
- DOMN/domaine par support de conservation
- 7 classes:
  - dessin
  - estampe
  - imprimé
  - miniature
  - peinture
  - photographie
  - sculpture

- 911 images per class
- Best top-1 test accuracy – 83%

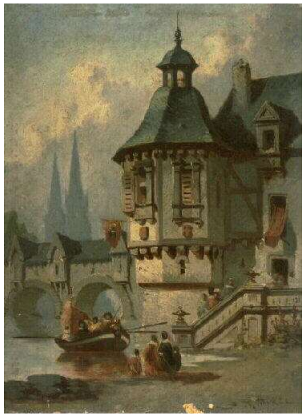


Classifying Art Forms

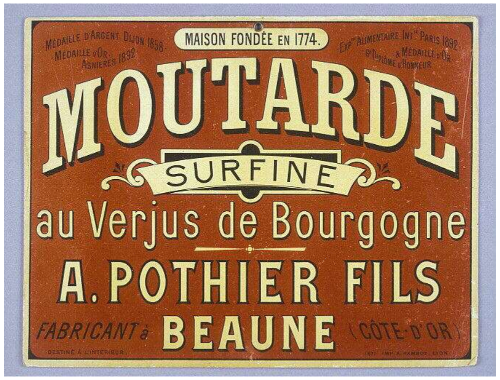
- Correct classification
  - Target: sculpture
  - Prediction: sculpture
    with prediction score 97.92%



  - Lable: peinture
  - Prediction: peinture
    with prediction score 73.73%

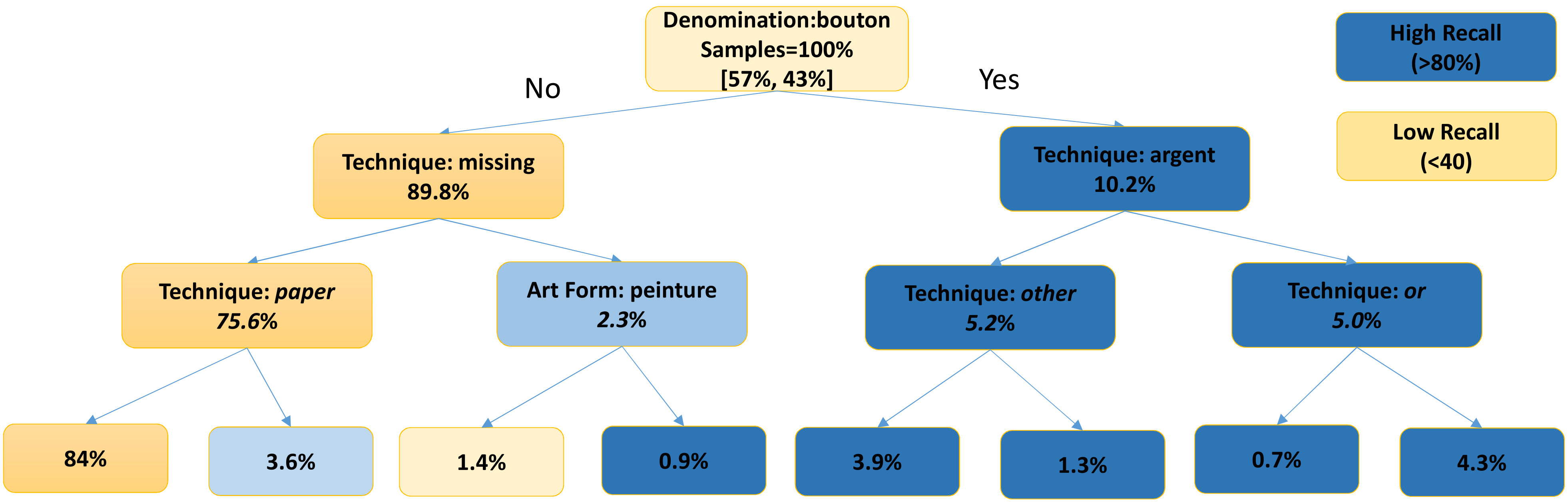


- Incorrect classification (maybe the case of mislabeling?)
  - Label: estampe
  - Prediction: imprimé
    with prediction score 99.72%



Statistical Analysis

- The classification outcome analysis shows some dependency on the image properties such as image aspect ratio and artwork properties such as art form, art media and technique, photographer and a museum.



Decision tree for the Animals classification with 6 statistically significant dependent variables.

Conclusions

- Deep Learning Transfer works even from non cultural collections
  - The model gives good accuracy results even on the classes that it hadn’t been pre-trained on.
- The limitations are mostly from the labeled image availability.
  - Adding images from other museum collections and Wikipedia can increase the training set.
- Joconde database metadata can be improved by filling the missing values and structuring some fields
  - Missing values can be inferred by machine learning algorithms.
  - Introducing the ontologies on techniques, material, denomination, preservation state can help to deeper explore the relations between all possible structured representations

Next Steps

Château de Bouillon : tête d'escalier by JEAN-HAFFEN Yvonne

- With the knowledge of the Joconde dataset and with the help of the subject matter experts define the set of classes of interest
- Build a model to recognize multiple object depicted by an artwork
- Improve the classification accuracy by performing various image transformations
- Augment the training data from the outside sources and/or by performing artistic transformations on ImageNet images
- Induction on RDF data and unstructured data

- **Explore combining:**
  - **Deep Network Layers and Thesaurus Layers (representation level)**
  - **learning and reasoning techniques (inference level)**

Thank you

- Acknowledgements:

Dr. Fabien Gandon

Dr. Frédéric Precioso

Liberté • Égalité • Fraternité
RÉPUBLIQUE FRANÇAISE

Ministère de la Culture