

DeepVisInterests : Deep data analysis for topics of interests prediction

Onsa Lazzez ¹, Adel Alimi ², and Wael Ouarda ²

¹Enis

²Affiliation not available

October 30, 2023

Abstract

Deep data analysis for latent information prediction has been an important research area. Many of the existing solutions have used the textual data and have obtained an accurate results for predicting users' interests and other latent attributes. However, little attention has been paid to visual data that is becoming increasingly popular in recent times. In this paper, we addresses the problem of discovering the attributed interest and of analyzing the performance of the automatic prediction using a comparison with the self assessed topics of interest (topics of interest provided by the user in a proposed questionnaire) based on data analysis techniques applied on the users visual data. We analyze the content of each user's images to aggregate the image-level users' interests distribution in order to obtain the user-level users' interest distribution. To do this, we employ the pretrained ImageNet convolutional neural networks architectures for the feature extraction step and to construct the ontology, as the users' interests model, in order to learn the semantic representation for the popular topics of interests defined by social networks (e.g., Facebook). Our experimental studies show that this analysis, on the most relevant features, enhances the performance of the prediction framework. In order to improve our framework's robustness and generalization with unknown users' profiles, we propose a novel database evaluation. Our proposed framework provided promising results which are competitive to state-of-the-art techniques with an accuracy of 0.80.

DeepVisInterests : Deep data analysis for topics of interests prediction

Onsa Lazzez^a, Wael Ouarda^c and Adel.M Alimi^{a,b}

^aREGIM-Lab.: REsearch Groups in Intelligent Machines, University of Sfax, BP 1173, Sfax, 3038, Tunisia

^bDepartment of Electrical and Electronic Engineering Science, Faculty of Engineering and the Built Environment, University of Johannesburg, South Africa.

^cDigital Research Center of Sfax, B.P. 275, Sakiet Ezzit, Sfax, 3021, Tunisia.

ARTICLE INFO

Keywords:

Deep data analysis
Convolutional Neural Networks
Online Social Networks
Deep learning
Ontology
Users' interest

ABSTRACT

Deep data analysis for latent information prediction has been an important research area. Many of the existing solutions have used the textual data and have obtained an accurate results for predicting users' interests and other latent attributes. However, little attention has been paid to visual data that is becoming increasingly popular in recent times. In this paper, we addresses the problem of discovering the attributed interest and of analyzing the performance of the automatic prediction using a comparison with the self assessed topics of interest (topics of interest provided by the user in a proposed questionnaire) based on data analysis techniques applied on the users visual data. We analyze the content of each user's images to aggregate the image-level users' interests distribution in order to obtain the user-level users' interest distribution. To do this, we employ the pretrained ImageNet convolutional neural networks architectures for the feature extraction step and to construct the ontology, as the users' interests model, in order to learn the semantic representation for the popular topics of interests defined by social networks (e.g., Facebook). Our experimental studies show that this analysis, on the most relevant features, enhances the performance of the prediction framework. In order to improve our framework's robustness and generalization with unknown users' profiles, we propose a novel database evaluation. Our proposed framework provided promising results which are competitive to state-of-the-art techniques with an accuracy of 0.80.

1. Introduction

The last decades have witnessed the boom of deep data analysis techniques with the huge amount of textual and visual data provided in many social networks (e.g., Facebook). Data analysis system provide a useful and efficient way to power prediction-required systems[3]. The great advancements on these systems using deep learning and its improved accuracy abilities, compared to traditional machine learning methods, have led to renewed efforts in social network analysis [6]. Hence, managing and understanding the data provided in social networks (e.g., Facebook), are still important research challenges [20] that are useful for diverse applications.

According to this, we focused our work on the data analysis techniques by the combination between the convolutional neural network architectures and the knowledge semantic representation to model the users profile with the ontology and to predict the user's interests distribution. Particularly, we focus on pretrained ImageNet Convolutional Neural Networks architectures (CNN) [11], which have become a highly recommended feature descriptor in many computer vision areas [4].

However, our proposed framework, *DeepVisInterest*, performs the users' interest prediction task based on a deep neural approach for the ontology construction and a list of topics of interest illustrated in table 1.

We increased the classification accuracy, in addition to the ability to handle sophisticated presentation attack conditions. Consequently, this would greatly hinder the effectiveness of users' interest prediction systems.

The major contributions of this paper consist of:

- Developing a novel framework named DeepVisInterests that performs the users' interests prediction task based on the pretrained ImageNet convolutional neural networks.
- Designing a new ontology using a set of deep visual features in order to learn the semantic representation for the popular topics of interests.
- Evaluating the system's generalization ability by conducting our novel database testing.

This paper is structured as follows: First of all, we start by reviewing the relevant works in section 2. Then, we provide our proposed database in section 3. Next, we elucidate a comprehensive explanation of our proposed approach in section 4 giving details pertaining to each phase. Finally, section 5 includes the achieved experiments with obtained results. In section 6 we discuss the achieved results on our constructed database. Meanwhile, we enclose our paper with some notes in section 7.

2. Relevant Works

The literature provides a various methods and purposes of deep data analysis. Due to the rapid advancement of deep learning, the data analysis has become more challenging to discover various users' latent attributes in practically the users' interests . In fact, the users' interest discovery process require the modeling phase to enhance the predicting phase.

ORCID(s):

Table 1

List of 24 topics of interest in social networks (e.g., Facebook)

TOI1: Sport	TOI2: Outdoors	TOI3: Food	TOI4: Drink	TOI5: Business	TOI6: Industry
TOI7: Travel	TOI8: Places	TOI9: Events	TOI10: Shopping	TOI11: Fashion	TOI12: People
TOI13: Activities	TOI14: News	TOI15: Entertainment	TOI16: Family	TOI17: Relationship	TOI18: Education
TOI19: Technology	TOI20: Lifestyle	TOI21: Culture	TOI22: Fitness	TOI23: Wellness	TOI24: Hobbies

2.1. Users' interest modeling

In [15], the authors suggested a Latent Topic of the users' interest (LUI) model to manage the topics distribution of tweets which possesses non-Gaussian characteristics. To evaluate their model, the authors employed two microblogs Weibo and Twitter to construct two dataset that contain respectively 10 million tweets and 100 million tweets and ultimately obtained a correlation coefficient between topics of 64.0% and 63.0% for each mentioned dataset.

The work proposed on [5] presented a users' interest model based on Latent Dirichlet Allocation to model topics from forums so as to distinguish between the users' interest topics and the unserious interest topics. To validate their model, the authors employed a forum thread from Tianya nad they obtained an accuracy of 80.5% and 93.3% for serious and unserious users respectively.

Also, Yang et al. [22] proposed a framework that integrates users' behavior on opinions and preferences aspects . The component has the capability to infer numerical ratings on the multiple aspects when such ratings are missing or not explicitly presented. To construct the users' model, firstly, they used LDA to cluster each aspect terms into latent aspects, secondly, the tensor factorization approach is applied to automatically extract weights of various aspects while approaching an overall numerical rating and finally a simple algorithm is used to computes an overall rating of an item based on both aspect ratings and weights. To validate their model, the authors used two real datasets: Internet Movie Database (IMDB) containing 193,266 reviews written by 83,585 users from the Internet Movie Database website and Hotel reviews Database containing 81,085 reviews from 879 users. In [17], the authors exploits the users' social data for developing aspects based sentiment analysis framework. This framework is based on the neighborhood CF algorithm, KL-divergence and a multidimensional Euclidean Distance to model the users' social data. To evaluate their proposed model, Musto et al., [17] used three database Yelp with 11,537 reviews from 45,981 users, TripAdvisor with 3,954 reviews from 536,932 users and Amazon with 50,210 reviews from 826,773 users.

2.2. Users' interest prediction

Some recent research efforts have been made to exploit data provided in social network using data analysis and deep learning techniques in users' interest prediction. In [24], the authors suggested a novel approach for the image-level and group-level label propagation of the users' interest prediction. They employed the AlexNet architecture for deep vi-

sual feature extraction and the image-level similarity to propagate the label information between images in order to disseminate the topics of interest-level for all the user's images. To validate their approach, the authors used a novel database collected from Pinterest containing 6000 images from 300 users' account and obtained an accuracy of 43.0%.

In [21], the authors proposed a method of assessment following social users based suggestions on categorical classification interest. This method relies on the convolutional neural network architectures and on hierarchical topic of interest categorization. To validate their approach, the authors used a database containing 20,500 images collected from Pinterest and obtained an overall precision of 39.9%.

In [2], the authors presented a classification of social images using unsupervised learning algorithm for users' interest prediction. To verify the validity of the proposed method, the authors used a novel database containing 800 social images collected from Pinterest and obtained an accuracy of 68.0%.

Table 2 illustrates the topics of interest used by the already mentioned related works.

2.3. Discussion

Performing the users' interests prediction model needs deep understanding of a users' social data. From the literature it is obvious that, this prediction has been obtained by analysing explicit social data. Although, prior works have examined the performance of ontology and deep learning techniques for objects detection in mining users' latent information. In fact, obtaining more accurate prediction needs deep understanding of users' social visual data. Beginning with the users' interests modeling phase. The LDA approaches presents users' data with a multi-nominal distribution distribution of words and documents while TF-IDF considers the ratio of the frequency of terms in each word over the total number of terms as a topic of probability. These methods achieve better accuracy in various research areas but, one limitation within these methods is the representation of data employing the Bas-of-Words method.

Concerning the users' interest prediction, topic modeling method exploits the social data to predict latent information about social users such as the topics of interests. Plentiful research has been performed in the domain of users' latent information prediction. Hence, the main objective is to classify social data: textual or visual data into some classes, according to the prediction area (sentiments, opinions, preferences, etc) in order to understand the social users' behavior. Then, several attempts have been done to categorize the social data and the literature attain good performance but it suf-

Table 2

State of the art on Users' interest prediction: Topics of interest

Ref	Data Type	Topics of interest
[24]	Social images from Pinterest	TOI7(Cars and Motorcycles), TOI17(Celebrities), TOI15(Films and Movies and Book)
		TOI4(Food and Drink), TOI2(Gardening), TOI22(Health and Fitness) TOI9(Holiday and Events) TOI11(Men's Fashion, Women's Fashion), TOI19(Technology), TOI20(Humor), TOI1(Sport), TOI24(Animals and Architectures).
[21]	Social images from Pinterest	TOI5(Business), TOI14(News), TOI7(Travel), TOI18(Education), TOI11(Fashion)
		TOI19(Technology), TOI3(Food), TOI23(Health), TOI10(Technology), TOI9(Food) TOI23(Entertainment), TOI10(Shopping), TOI20(Hobbies), TOI24(Religion).
[2]	Social images from Pinterest	TOI15(Entertainment), TOI5(Business), TOI19(Technology), TOI23(Health), TOI14(News).

fer from some limits like the use of Facebook as the social data source instead of Pinterest. In fact, a very little works has been done that use the social visual data shared on Facebook for developing a supervised deep learning based users' interest prediction. Hence, Facebook present the most important social network to several users types and it becomes an important way to share a huge amount of data daily. Also, some recent effort have been made to exploit social data adopting deep learning approaches in this prediction and the mentioned works used the traditional topic modeling methods like LDA and feature extraction methods like AlexNet architecture.

3. VisualDatabase: Pictures and Interest

Data provided in social networks are considered as sensitive data as they reflect the private life of social media users. Therefore, creating a social database would require the approval of social media users to gather data so that their private life is respected. To reach this end, an application has been created [13] allowing users to voluntarily subscribe, and therefore, give us the permission to gather the data they publish on their social networks (e.g., Facebook) accounts/pages. Moreover, only abstract images with natural scenes and neutral themes have been selected to avoid personal images/pictures of the user's family or friends.

Our database, named VisualDatabase, is a set of multiple social images, which was created in March 2018. This database contains social data from 240 accounts, simply called social users. For each social user, there are 100 random images selected from the "liked" and "shared" ones. The database images have 320*320 resolution, with multiple ethnicity and locations (Africans) and from different age intervals (between 15 and 60 years old). Figure 1 shows some image samples. Furthermore, the database presents the self-assessed users' interest based on a big interest questionnaire (BI) voluntarily filled by each user. The result of the BI is a vector where each component indicates the disposition of a user with respect to the core topics presented in table 1. The self assessed traits are examined to be the validated user' score interest.

**Figure 1:** Image samples from VisualDatabase

4. Proposed Approach

The problem of users interest prediction may be considered as an image classification problem. However, in contrast to the traditional image classification where the objective is to maximize classification performance at individual level, we are based more on learning the overall user-level image distribution.

Our proposed framework, named *DeepVisInterest*, is illustrated in figure 2 and is based essentially on users' interest modeling phase followed by the prediction phase.

4.1. Phase 1: Users' interest modeling

The ontology have been conceived to allow for a common definition of concepts, entities, relationships, situations and events and consequently for common understanding and for promoting information exchange. In fact, we observe that ontologies have been used in deep learning techniques to address uncertainty in image classification based on ob-

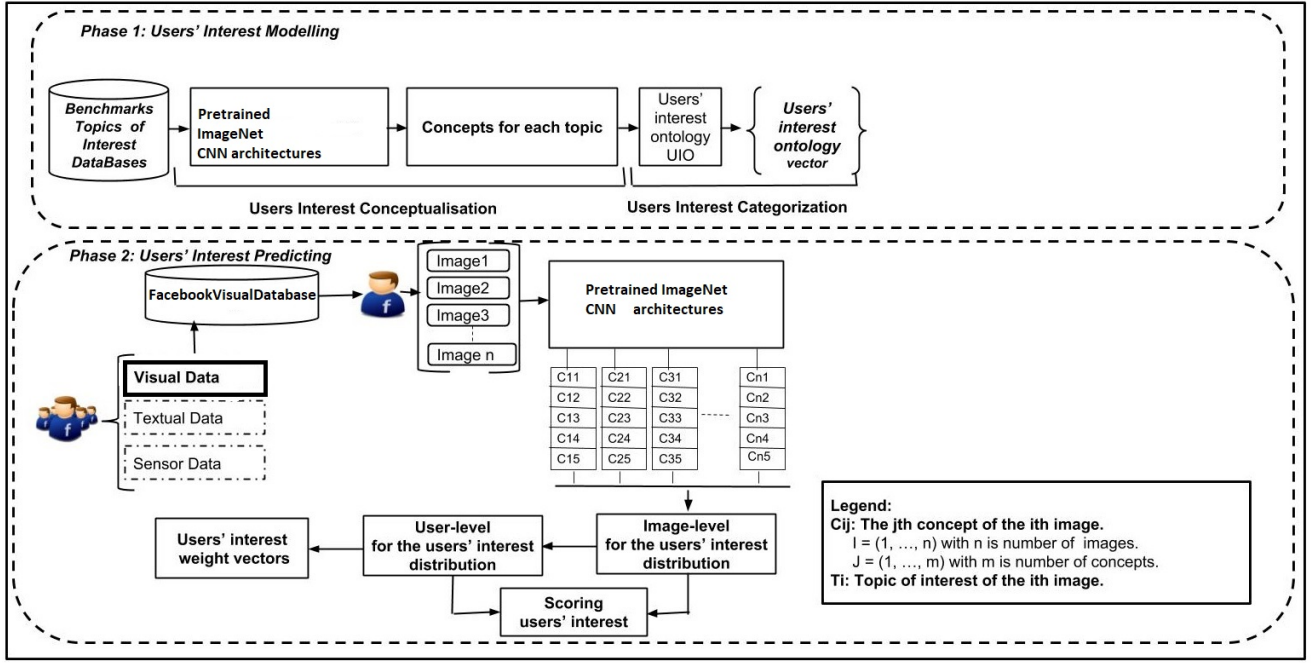


Figure 2: *DeepVisInterest*: Deep learning framework for users' interest discovery

ject detection methods [8]. This observation led us to exploit ontologies as a users' interest model to classify each user's images under 24 topics, already presented in table 1, and predict the user's topic of interest. An overview of our constructed ontology is illustrated in figure 3. In our proposed ontology, the sub-concepts are the 24 topics. We have used 24 benchmark database to exhibit the 24 topics (each database contains approximately 200 images). Taking an unannotated image as input, we employ the pretrained ImageNet CNN architecture for object detection to detect the top 5 ranked objects, among the 1000 objects of ImageNet, that will be incorporated into ontology as the end-concepts that are linked for the sub-concepts by the object property "is-a".

The main architecture comes from the CNN presented by Krizhevsky et al. [12] that has obtained the literature performance in the challenging ImageNet classification issue. For a more detailed description of these architecture, we direct the reader to the original paper. In fact, the ontology implementation features represent the number of concepts, data properties, object properties, individuals and axioms. To construct our ontology, we used the tool named Protégé versioned 4.3 and this also provides Resource Description Framework (RDF) schemas and XML scripts for using ontology in web.

The Protégé is an ontology development environment with a huge number of active users. This environment has been extended with support for OWL (Web Ontology Language) and has become one of the leading owl tools[1]. Our proposed ontology applies all the implementation features while the object properties are used to define the relationship between individuals.

Users' interest ontology vocabulary:

We defines the basic metrics for the size of the users' interest ontology on various aspects. The size of our ontology is defined as follows:

- $size_C(O) = ||C||$,
- $size_I(O) = ||I^c||$,
- $size_A(O) = \sum_{c \in C} ||A^c||$,
- $size_R(O) = \sum_{r \in R} ||r||$,
- $size(O) = size_C(O) + size_I(O) + size_A(O) + size_R(O)$

Let O be the users' interest ontology:

$Size_C(O) = 33$, $Size_I(O) = 5$, $Size_A(O) = 443$,
 $Size_R(O) = 32$ and $Size(O) = 1555$.

- I is a collection of finite sets indexed by C as follow
 $I = I^c | c \in C$.
- A is a collection of a set of attributes with
 $A = A^c | c \in C$. Each $\varphi \in A^c$ is an attribute of concept c .
- The value of each attribute φ for an instance $\alpha \in I^c$ of concept c , is presented by $\varphi(\alpha)$. $\varphi(\alpha)$ is either a data value or type T or an instance of concept c .
- R is a set of binary relations on the set of concepts.
 $R = r_1, r_2, \dots, r_k$. For each $r \in R$ and $r \subseteq C \times C'$,
 $(c, c') \in i$.

Users' interest ontology structure:

Structural metrics are the most immense examining metrics in the ontology presentation, exactly, cohesion metrics that measure the degree of relatedness between concepts. Among the cohesion metrics, we find the relation-based structural complexity. In fact, for each $r \in R$ we have some few structural metrics such as

- the number of Root nodes with
 $NRN_r(O) = ||Root_r(O)||$,
- the number of Leaf nodes with
 $NLN_r(O) = ||Leaf_r(O)||$,
- the maximum length of simple path with
 $MaxSPL_r(O) = Max_{p \in path(O)}(Lenght(P))$,
- the number of Isolated nodes as
 $NIC_r(O) = ||Root_r(O) \cap Leaf_r(O)||$,
- the total number of Reachable nodes from Roots with
 $TNRNR_r(O) = \sum_{x \in Root_r(O)} ||Reachable_r(O)||$ and
- the average number of Reachable nodes from Roots with
 $ANRNR_r(O) = TNRNR_r(O) \setminus ||NRN_r(O)||$.

For the users interests ontology, the "is-a" relation based structure metrics are:

- $NRN = 1, NLN = 1, MaxSPL = 3$ and
- $NIC = 0, TNRNT = 36, ANRNR = 36$.

Users' interest ontology context:

We focus on users' interest predicting and is interested in if an ontology is a perfect tool for modeling the semantics of topics of interests.

Let assume that a user U possesses n topics of interests $UI_1, \dots, UI_n$ which contain a set of concepts $C_{i,1}, \dots, C_{i,n}$ and a set of attributes $A_{j,1}, \dots, A_{j,k}$.

An ontological description of the semantics of such a topic of interest consists of the following expressions:

- EXP_{UI_i} which describes the functionality of the topic UI_i ,
- $EXP_{C_{i,j}}$ which defines the meaning of the parameter $C_{i,j}$ in the set of concepts,
- $EXP_{Y_{i,k}}$ which illustrates the meaning of the parameter $Y_{i,k}$, in the set of attributes of each concept.

Semiotic metrics assessment:

The quality of each ontology is defined across a set of semiotic metrics. These metrics assess the syntactic, semantic, pragmatic and social aspects of ontology quality. Then, we use Protégé-OWL-5.2 which is an open-source platform that provides tools to construct domain models and knowledge-based applications with ontology.

Figure 3 illustrates the incorporation of CNN architectures

outputs in our ontology. Each input image is represented by a vector containing a set of scores of the top 5 concepts among 1000 concepts of ImageNet [7] then the ontology vectorization method is used to find each user's interest using OWL API.

4.2. Benchmarks

To model the users' interest, we conducted our ontology construction on twenty four publicly available databases, that represent the twenty four topics mentioned in tabel 1, such as: Food Images database [18], Sport Event Database [14] and DeepFashion Database [16].

These are challenging databases, since they consist of various types of users' behavior with different image qualities (high definition, average, and very low quality). In the following, we give a detailed description of these databases.

4.2.1. Food Images DataBase

The database comprises 568 food images including sweet (e.g., ice cream, chocolate), savory (e.g., pistachios, sandwiches), processed (e.g., hamburger, French fries, potato chips, chocolate bars) and whole foods (e.g., vegetables and fruits) and beverages (e.g., coffee, orange juice). Images were selected from a commercially available database (Hemera Photo Objects). All images are color photographs with a resolution of 600×450 pixels.

4.2.2. Sport Event DataBase

The Sport Event database contains 600 images including 6 sport event categories that are rowing, badminton, polo, bocce, snowboarding, croquet, sailing, and rock climbing. Images are divided into easy and medium according to the human subject judgement. Each image provides some information of the distance of the foreground objects. All images are color photographs with a resolution of 600×450 pixels and they are collected from non-copyrighted sources on the internet.

4.2.3. DeepFashion DataBase

The DeepFashionDB contains 800 diverse fashion images ranging from well-posed shop images to unconstrained consumer photos. All images are color photographs with a resolution of 600×450 pixels.

4.3. Phase 2: Users' interest predicting

Our prediction phase requires the construction of Visual-Database to propagate the topics of interest distribution from image-level to user-level.

4.3.1. Visual Users' interest prediction

The visual users' interest prediction is based on the combination between the image and user-level (VUIP-IL/UL) methods. Figure 4 illustrates the main steps of our proposed method named VUIP-IL/UL. This method is based on 3 principal steps. In the first step, we apply the pretrained ImageNet CNN architectures for object detection to extract the deep visual features from VisualDatabase. These features display the image objects with their probabilities. Thus, by

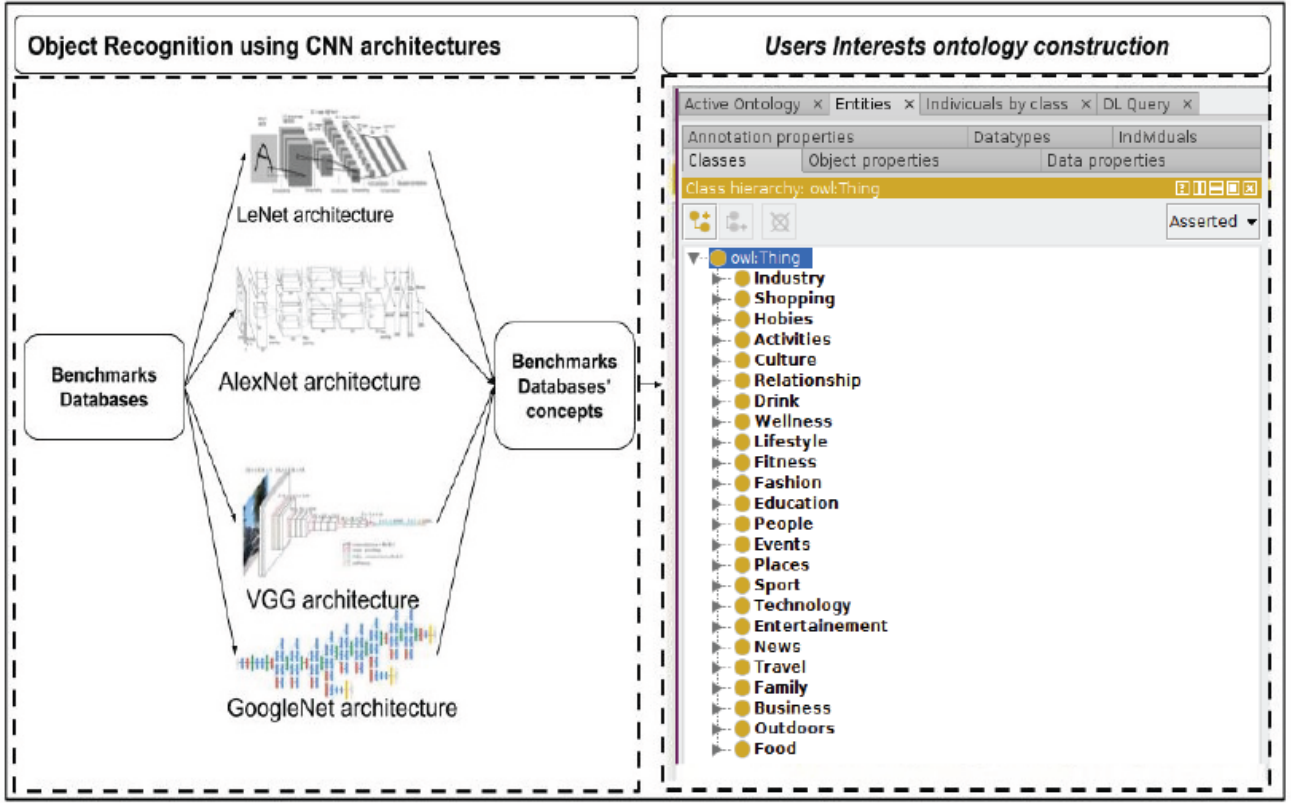


Figure 3: Users' interest ontology construction based on CNN architectures

inferring our users' interest ontology (UIO) model, each image object is replaced by its corresponding super-concept. In the second step, the probability and occurrence-based scoring mechanisms are applied to obtain the image-level distribution. In the third step, we build a mapping matrix from this level to that of the user-level.

Scoring users' interest: To quantify the users' interest, once the user's interest are predicted, we use a scoring function to find the weight of each topic. This mechanism is very powerful as the user's interest scores will be applied to determine the adapted interest distribution for each user's image and therefore, for each user. We use a probability and occurrence-based scoring mechanisms. The topic's score of each image $i \in I$ posted by a given user $u \in U$, may be measured by the probability and occurrence of an object $o \in O$ where an image is represented by a collection of objects O .

$$S(u, I) = \sum_{i=1}^n p_{oi} oc_{oi} \quad (1)$$

where $p_{oi} \in P_I$, $oc_{oi} \in OC_I$.

Here, P_I is a set of probabilities within each $i \in I$ obtained by each image' object and OC_I is a set of occurrences with object $o \in O$ for the given image. Algorithm 1 demonstrates the detailed steps used in our prediction task.

- BD: List of 24 benchmark databases presented the 24 core topics of interest.

- VisualDatabase: Test database.
- P: Pretrained CNN models.
- U: Users in test database.
- I: collection of shared images of each $u \in U$.

1. Extract the deep visual features from P of BD.
2. Users' interest ontology construction: UIO.
3. UIO vectorisation.
4. Extract the deep visual features from P of I.
5. Image-level for the users' interest distribution.
6. Return G, G': weight of the matrices specific for I of every $u \in U$.
7. User-level for the users' interest distribution.
8. Return V: weight vector specific for each $u \in U$.

4.3.2. Image-level for the users' interest distribution

After applying the feature extraction step, each image possesses 5 objects with their probabilities; such as (espresso, 0.08), (cup, 0.07), (dough, 0.06), (ladle, 0.05) and (sandal, 0.04). Using the Fact++ reasoner and DL query, we infer the users' interest ontology to result the super-class for each image object. We use the data property "has-Instance" in order to generate in the super-class for each input image's object presented as an ontology instance. This step applies the Fact++ reasoner and the DL query: (has-Instance value "image object"). For example, (has-Instance value "espresso")

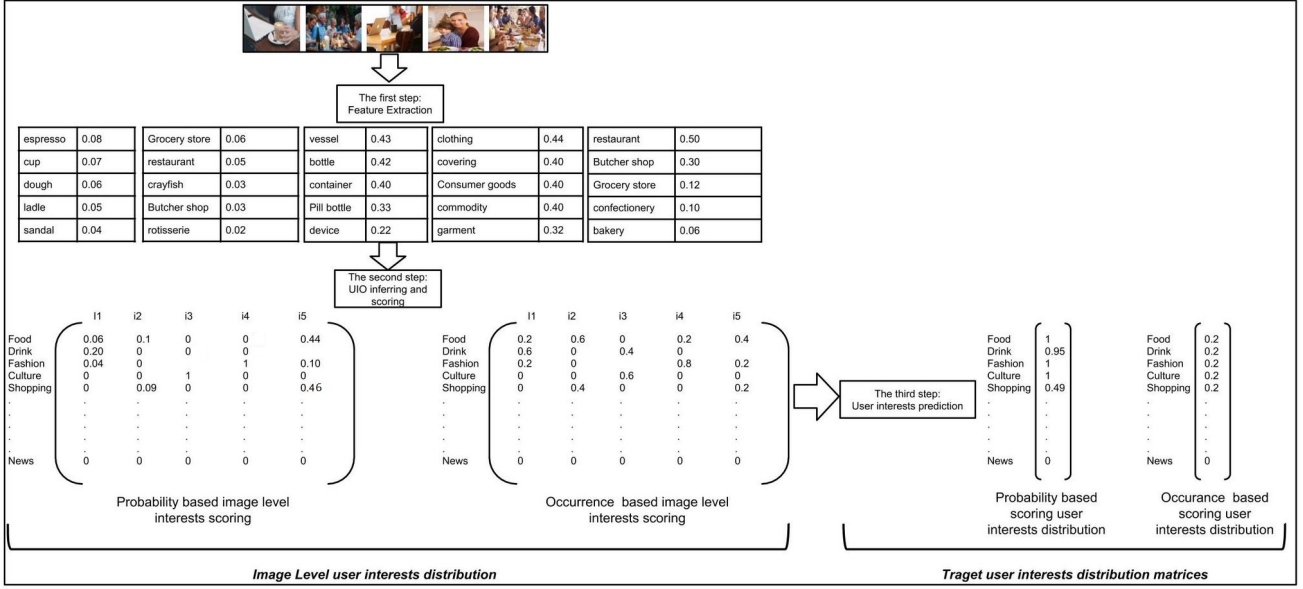


Figure 4: An illustration of VUIP-IL/UL method. Three steps are incorporated.

results in the super-class "Drink".

For the image topics scoring, we apply two mechanisms;

- For the probability based scoring, we define $S(i, t) = \sum P(o, i, t)$ where $S(i, t)$ is the score of image i in the topic t with $t \in T$ and $P(o, i, t)$ is the probability of object o of image i with $o \in O$ and O is the set of image' objects with their super class in UIO are t .
- For the occurrence based scoring, $S(i, t) = \frac{N(o, i, t)}{5}$ where $N(o, i, t)$ is the number of image' objects with their super class in UIO are t .

For more details, see the Algorithm 2. Accordingly, we describe two matrices $G \in R^{n \times 24}$ and $G' \in R^{n \times 24}$ to be the affinity matrices between the twenty-four core topics of interest and the n shared images by a specific user as $u \in U$.

- U : users in the test database.
- UIO : users' interest ontology.
- T : List of 24 core topics of interest.
- $I = i_1, i_2, \dots, i_n$: n shared images by specific $u \in U$
- P : Pretrained ImageNet CNN models.

1. Apply P for object recognition on each $i \in I$ to extract the five objects with high probability.
2. Infer UIO , using Fact++ reasoner and DL queries to predict the super class for every object per $i \in I$.
3. Scoring image-level of the users' interest distribution.
4. Repeat :
5. Employ 1, 2 and 3 for all I .

6. Until n for the three cases: $n = 5$, $n = 10$, $n = 50$, $n = 75$ and $n = 100$.

7. Return G , G' : weight matrices of probabilities and occurrence based scoring mechanism.

4.3.3. User-level for the users' interest distribution

According to the already-explained image-level, each user $u \in U$ possesses two weighted matrices G and G' for n shared images on social networks (e.g., Facebook). For this level, we intend to generate the target user's interest distribution matrix based on the two scoring mechanisms which are (a) it first treats the matrix G , we define:

$$S(u, t_k) = \sum_{k=1}^{24} \sum_{i=1}^n p_{i,k} \quad (2)$$

where $S(u, t_k)$ is the user's score u about the topic t with $p_{i,k}$ is the probability of image i for topic k and n is the number of shared images.

(b) the second mechanism treats the matrix G' , we define for $k=(1,24)$ and $i=(1, n)$:

$$S(u, t_k) = \frac{N(\max(p_{i,k}))}{n} \quad (3)$$

where $S(u, t_k)$ is the user's score u about topic t with $p_{i,k}$ is the probability of image i for topic k and n is the number of shared images. For more details, see the Algorithm 3.

[!htb] [1]

- U : users in the test database.
- $I = i_1, i_2, \dots, i_n$: n shared images by specific $u \in U$
- G , G' : weight matrices of probabilities and occurrence based scoring mechanism respectively.

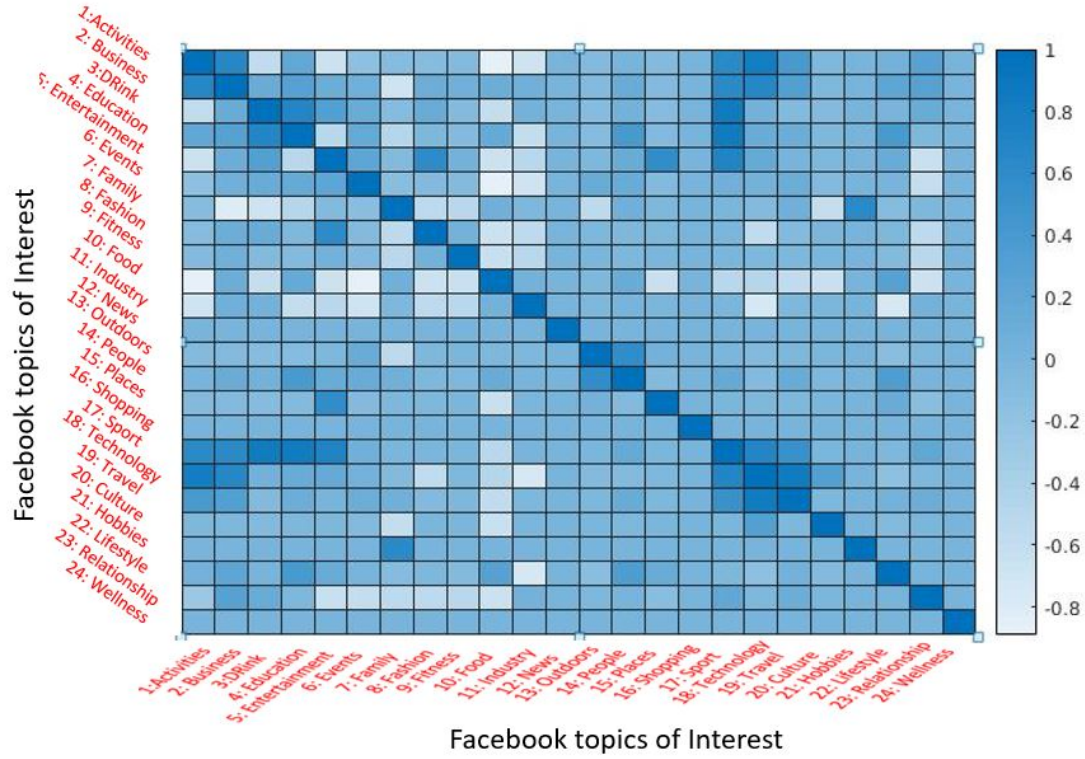


Figure 5: Pearson Moment Coefficients between various topics of interest in Table I.

1. Extract target user's interest distribution vector V for probability based scoring for the three cases: $n = 5$, $n = 10$, $n = 50$, $n = 75$ and $n = 100$.
2. Extract target user's interest distributions on vector V' for occurrence based scoring for the three cases: $n = 5$, $n = 10$, $n = 50$, $n = 75$ and $n = 100$.
3. Return V, V' .

5. Experimental Analysis and Discussion

In this study, our aim is to investigate the impact of visual data on users' interest prediction problem. To this end, we led an intensive experimental study by generating multiple models from different images combinations. We have eventually reported the obtained results from each combination. We use the publicly available implementation Caffe [9] to test our model. All of our experiments are evaluated on a Linux X86-64 machine with 32 GRAM.

5.1. Correlational study between topics of interest

Figure 5 illustrate the dependency between various topics already mentioned in table 1 using the Pearson Moment correlation.

This coefficient is commonly represented by $p(rho)$ with:

$$p_{x,y} = \frac{Cov(X, Y)}{\sigma_x \cdot \sigma_y} \quad (4)$$

where Cov is the covariance, σ_x is the standard deviation of X and σ_y is the standard deviation of Y .

As indicated in figure 6, firstly, we notice that the topic Food is totally highly positive correlated with the topics Drink with a value greater than 0.5. This high correlation means that the images containing objects belonging to the super-class Food, in our UIO ontology, may contain objects belonging to the super-class Drink or Family or People. Secondly, we observe that the topic Fashion is highly negative correlated with topics Technology with a value -0.553 which means that the user who has Fashion as a topic of interest can never be interested in topics Technology. Finally, the topic Education, for example, is correlated with the topic Culture with a value 0.033.

5.2. Quantitative study of the users' interest prediction

5.2.1. Image-level for the users' interest distribution study

To illustrate this level, we propose a demonstration about the *Family* class. According to the figure 7, we remark that this distribution is articulated around three topics which are *Outdoors*, *Drink*, *People* and *Food*. These results, validate the positive correlation between these topics and the fact that a user who is interested by the topic *Family*, he/she may

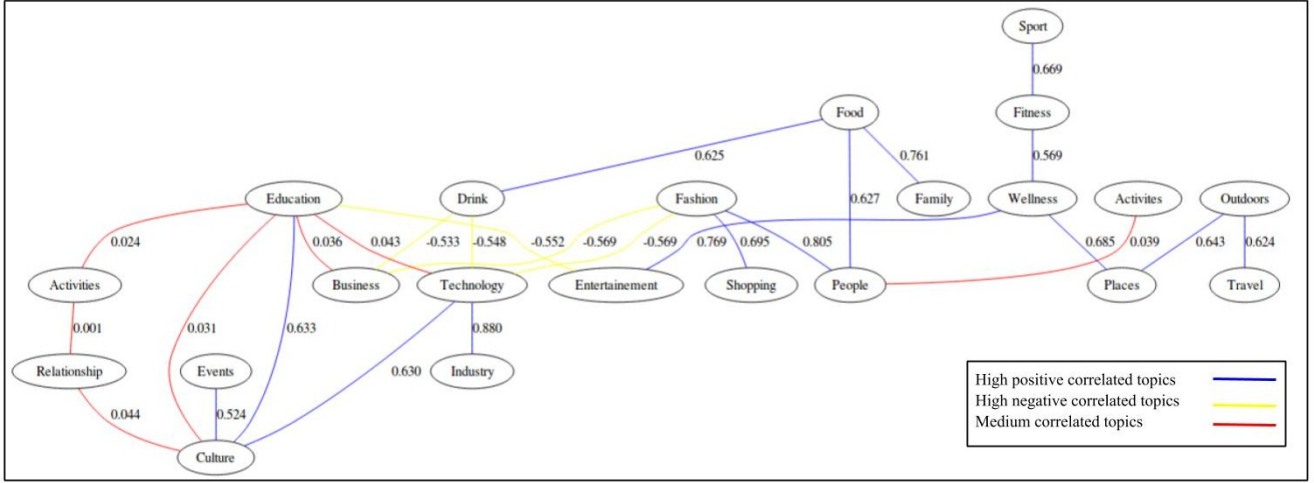


Figure 6: Topic of interests correlation

share images belong to these discovered topics. For the same class, the figure 8 illustrates this distribution level using 100 shared images per user. In fact, with 100 images, we remark that the images distribution has become more detailed with the appearance of new topics with low scores such as *Shopping*, *Places* and *Entertainment* and the disappearance of the self-assessed topic for some users. This appearance and disappearance is explained by the diversity of images that can generate vectors with low scores for several topics.

To conclude, we can assume that in the social networks (e.g., Facebook), each user may share a set of images whose can be related or not related to his/her self-assessed topic of interest. With 50 images, the distribution is generally very close to reality with the appearance of the self-assessed with the most important score and with 100 images, this distribution has become more detailed with new topics assigned to that are correlated with self-assessed topic for each user.

5.2.2. User-level for the users' interest distribution study

At this level, we attempt to define the confidentiality area of the users' shared images. For this reason, we used k shared images by each user within the 24 classes. From each class, we obtained a confidentiality area which generates the target users' interest matrix with a high score for the self-assessed topic. Table 5 describes the variation of the accuracy measure for each class according to the number of shared images per user. This variation is assigned to the fact that each user's topics of interests distribution consists of three layers: starting term with 5 images, middle term with 10 images, long term with 50 images, very long term with 75 images and extreme term with 100 images.

- The starting term presents the sharing of the first 5 images that each user chooses indicating their self-assessed topic of interest,
- For the middle term, the same user may be influenced

by other topics and can share some images that can interrupt our classification, which explains the decrease of the system performance from 0.85 to 0.75,

- In the long term, after being biased in the middle term, the user settles back in his/her self-assessed topic of interest while our system predicts the correct target class with 50 shared images to obtain an accuracy of 0.95,
- In the very long term, the user keeps a stability with a slight disturbance of his/her distribution obtained in the long term. For this reason, our system shows a slight decrease of performance with an accuracy of 0.80,
- In the extreme long term, our system performance undergoes a very remarkable decrease with an accuracy of 0.65, which validates that beyond 50 images, the distribution of the topics of interest encounters a disturbance by the diversity of the images, which negatively influences the self-assessed topic.

To better visualize this variation, figures 9-12 describe the Cumulative Match Characteristics (CMC) curves. We remark that some classes possess a high accuracy value for a specific number of images. Other classes present the stability of accuracy value some number of images. In addition, divers classes possess an accuracy of 0 for some number of images. For example, Lifestyle class presents an accuracy of 0 for 5, 10, 50, 75, 100 shared images and News class has an accuracy of 0 for 5, 10, 50 images. In fact, the users who have Lifestyle or News as self-assessed class are more likely to have perturbation in their shared images that change the output class to any other target class. In this context, we try to discuss the reasons why some images are misclassified over classes other than the self-assessed class through the figure 13.

To evaluate our framework, we deliver a continuously growing set of pre-trained models with famous architectures

Table 3The *DeepVisInterest* accuracy measure variation of

Class	Images' number				
	5	10	50	75	100
Activities	0.3	0.3	0.3	0	0
Business	0.7	0.7	0.7	0.4	0
Drink	0.6	1	0.9	0.4	0.4
Education	1	1	1	0.4	0
Entertainment	0.9	0.6	1	0.7	0
Events	1	1	0.8	0.3	0.3
Family	0.4	0.8	1	0.6	0
Fashion	1	1	1	0.7	0
Fitness	0.4	0	1	0.3	0
Food	1	1	0.9	0.6	0
Industry	1	0.6	0	0.7	0
News	0	1	0	0.5	0
Outdoors	1	0.5	1	0.6	0
People	0.4	1	0.7	0.2	0
Places	0.7	1	0.4	0.5	0
Shopping	0.5	0.2	0.5	0.5	0
Sport	0.5	0.4	0.5	0.5	0
Technology	0.6	1	0.7	0.5	0
Travel	1	0.2	0.4	0.3	0
Culture	0.6	0.4	1	0.3	0
Hobbies	0.5	1	1	0.3	0
Lifestyle	0	0	0	0.2	0
Relationship	1	0	0.7	0.8	0
Wellness	0	0	0.5	0	0

Table 4The variation of *DeepVisInterest* performance

Number of shared images per user	Accuracy
5 images	0.85
10 images	0.75
50 images	0.95
75 images	0.80
100 images	0.65

for the Caffe framework [10]. One focus in our work is the depth of the CNN, which affects the capability of the convolutional layers. Thus, we make use of four different CNN architectures illustrated in table 7 in terms of test accuracy performance. For that, GoogleNet architecture is the best architecture based on the idea of executing the layers in parallel with the inception module. It helps us to get a better classification accuracy by extracting the information about the very fine grain details in the depth.

To highlight our framework, from table 8, we can assume that our framework outperforms the literature meth-

Table 5The *DeepVisInterest* accuracy measure according to features extraction method

Pretrained CNN Architecture	Accuracy
<i>DeepVisInterest</i> with LeNet	0,760
<i>DeepVisInterest</i> with AlexNet	0,810
<i>DeepVisInterest</i> with VGG19	0,910
<i>DeepVisInterest</i> with GoogleNet	0,930

ods on several levels. We can valorize our work by choosing Facebook as a source of data among other social networks through the implementation of a specific Facebook application [13] compared with other works [19, 24, 23] that use Pinterest as social source applying the crawling method based on a public APIs. In addition, we apply four CNN architectures for object recognition to enhance the feature extraction and classification modules compared to those of other works. To evaluate the performance of our algorithm, we apply two diverse criteria **1) Precision** and **2) Recall** illustrated in table 9 and table 10 respectively.

We used the Cumulative Match Characteristics (CMC) curve to illustrate the evolution of interests prediction rate with the number of shared images and to compare this rate for each class.

6. Result's Discussion

In this section we will discuss our obtained results by evaluating the performance of each architecture used for features extraction phase. Each CNN architecture contains two separate modules that are the feature extractor and the classification modules. The performance of any CNN architecture is related to the feature extractor module parameters, especially the number and size of filters and layers.

AlexNet is one of the deep convolutional neural network to deal with complex scene classification task. AlexNet has 5 convolutional layers, 3 sub sampling layers and 3 fully connected layers. It use a set of filters with size of $11 * 11$, $5 * 5$ and $3 * 3$ respectively for each convolutional layer.

Furthermore, to achieve better performance, the complexity of convolutional neural networks is continually increasing with deeper architectures. The VGG'19 is much deeper than AlexNet with 19 layers including thirteen convolutional layers with filters size of $3 * 3$ and 3 fully-connected layers. The use of very small filters sizes capture a set of fine deep visual features from the image input, decrease the parameters number and increase the filters number. The increase of filters number augment the depth of the input image and consequently the depth of the network which presents a critical component for good performance in the users interests classification task based on the complex scene image classification.

Given that VGG'19 is based on the filters simplicity and the depth of the network, GoogleNet architecture is one of the first architecture that introduces the idea of executing the layers in parallel with the inception module based on the concatenation operation at different scale.

Table 6
Comparison between *DeepVisInterests* and State-of-the art

Approach	Social Network	Number of images	Feature extraction	Number of classes	Accuracy
[24]	Pinterest	300	AlexNet	7	0,430
[19]	Pinterest	800	AlexNet and VGG19	5	0,680
[23]	Pinterest	20500	Siamese	2	0,399
<i>DeepVisInterest</i>	Facebook	24000	CNN based Ontology inference	24	0,850

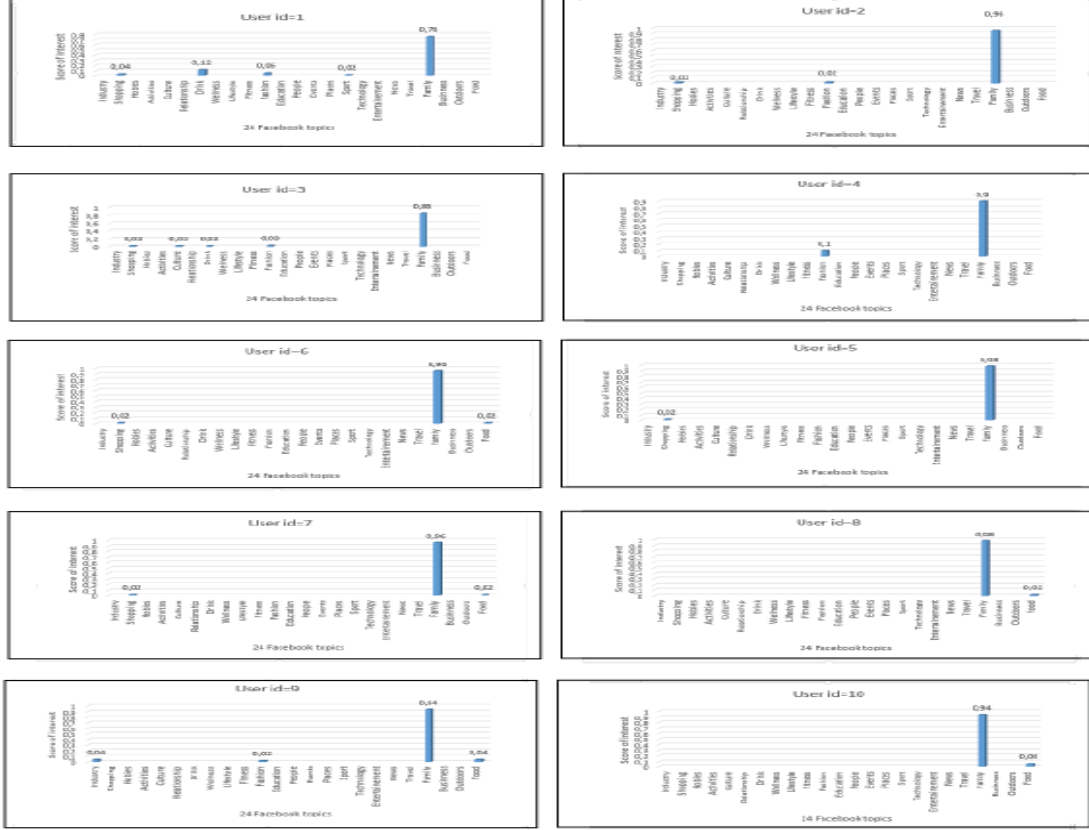


Figure 7: Image-Level distribution for Family class with 50 images per user.

GoogleNet use 9 inception modules with a creative structuring of layers in order to improve performance and computationally efficiency. Hence, GoogleNet help us to get better classification accuracy by extracting information about the very fine grain details in the volume.

After having discussed the efficiency of feature extraction phase, we try to analyse the images that are misclassified over classes other than the self-assessed class. This misclassification is caused by divers adversarial attacks in the form of delicate perturbations to each input user' image that conduct our framework to predict incorrect class compared to the self-assessed class. However, the users which have self-assessed classes like *Relationship*, *News*, *Wellness*, *Lifestyle* and *Hobbies* are more likely to have perturbations in their shared images that change the output class to any other tar-

get class.

Several examples of miss-classified users are shown in figure 13. The labels above the selected images present the self-assessed classes for each group of users and the labels below the images are the target classes obtained by our framework. In the first group on the right, we see a set of images shared by users, whose have *News* as self-assessed class. The classification of these images is based on the objects belong each one obtained by CNN architecture for objects recognition. These objects possess *Fashion* or *Places* or **People** as super class in our UIO ontology. This inference presents a perturbation to predict other target classes other than the self-assessed by the user.

In the second group in the middle, we illustrate some images shared by users whose have *Relationship* as self-assessed

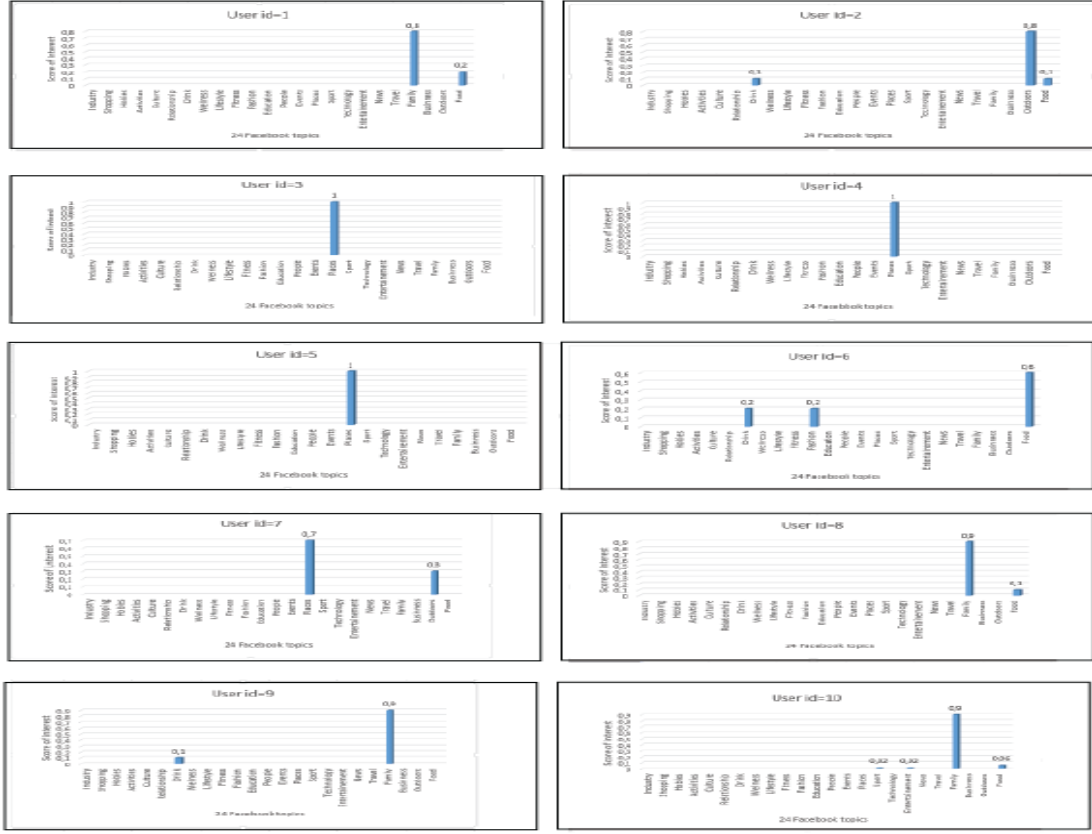


Figure 8: Image-Level distribution for Family class with 100 images per user.

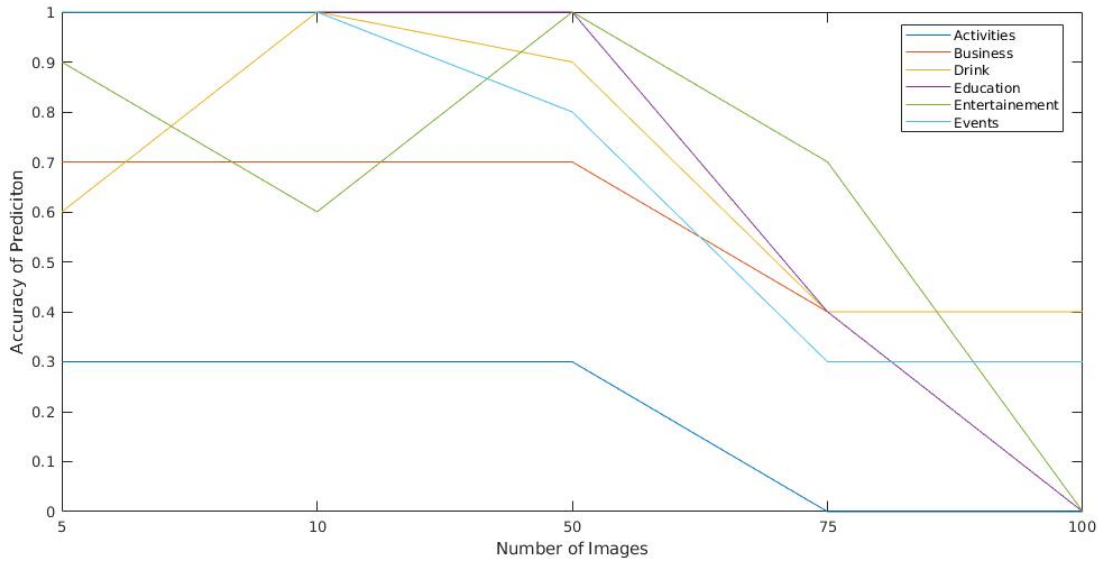


Figure 9: Cumulative Match Characteristics (CMC) curve for Activities, Business, Drink, Education, Entertainment and Events classes

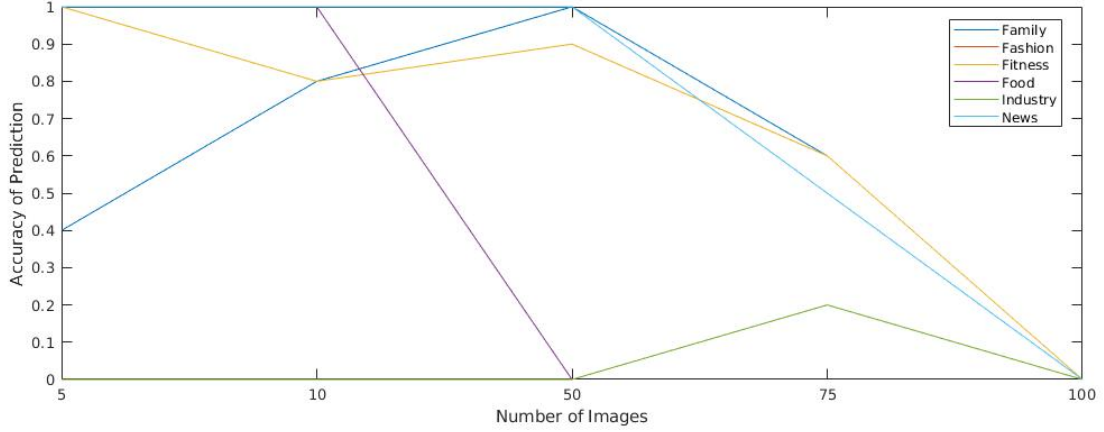


Figure 10: Cumulative Match Characteristics (CMC) curve for Family, Fashion, Fitness, Food, Industry and News classes

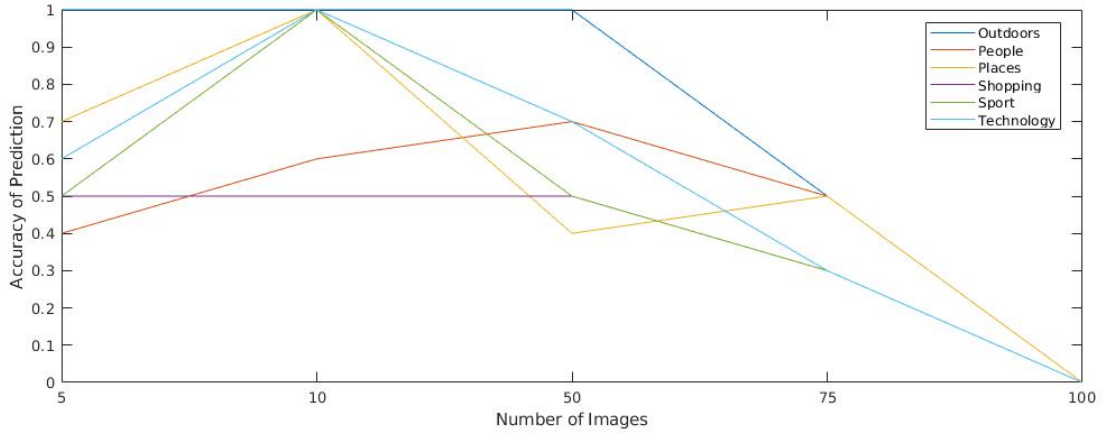


Figure 11: Cumulative Match Characteristics (CMC) curve for Outdoors, People, Places, Shopping, Sport, Technology classes

class. The classification of these images product *Outdoors* and *People* as target classes. This miss-classified is caused by the fact that the objects belong these shared images have *Outdoors* or *People* as super class in our UIO ontology. In fact, the topic *Relationship* reflects the importance of making friendships with people in order to have a good time together.

In the third group on the left, we present divers images shared by users whose have *Wellness* as self-assessed topic. The fact that these shared images contains objects which have *Outdoors* or *People*, these users have *Outdoors* or *People* as target classes in our classification.

7. Conclusion

In this paper, we have examined the correlation between topics of interest in social networks (e.g., in Facebook) and the definition of the confidential area of number of shared images to obtain the best user's interest distribution. A joint novel framework, named *DeepVisInterest*, was established to predict users' interests from visual data applying mainly

the CNN architectures for feature extractor and classification modules. We have introduced novel users' interest model to conceptualize and categorize the 24 topics of interests into semantic representation using the ontology. We have systematically evaluated the proposed framework regarding our VisualDatabase which contains over 24000 images. Our system has shown competitive results compared to other state of the art techniques.

Acknowledgements

The research leading to these results has received funding from the Ministry of Higher Education and Scientific Research of Tunisia under the grant agreement number LR11ES48.

References

- [1] , 2018. Assessing the quality of domain ontologies: Metrics and an automated ranking system. *Data Knowledge Engineering* .
- [2] , 2020. Cnn-based malicious user detection in social networks. *Concurrency and Computation: Practice and Experience* 30.

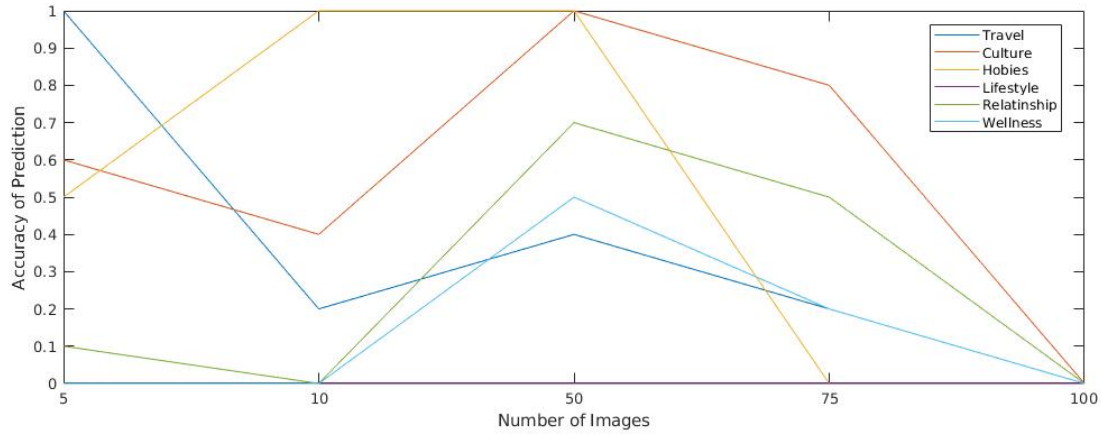


Figure 12: Cumulative Match Characteristics (CMC) curve for Outdoors, People, Places, Shopping, Sport, Technology classes



Figure 13: Some adversarial examples of images: (a) Users who have "News" as a self-assessed class, (b) Users who have "Relationships" as a self-assessed class and (c) Users who have "Wellness" as a self-assessed class

- [3] Alam, F., Alam, T., Offi, F., Imran, M., 2021. Social media images classification models for real-time disaster response. CoRR abs/2104.04184. URL: <https://arxiv.org/abs/2104.04184>, arXiv:2104.04184.
- [4] Alpaydin, E., 2020. Introduction to machine learning. MIT press.
- [5] Chen, C., Ren, J., 2017. Forum latent dirichlet allocation for user interest discovery. Knowledge-Based Systems 126, 1 – 7.
- [6] Chung, T.L.D., Johnson, O., Hall-Phillips, A., Kim, K., 2021. The effects of offline events on online connective actions: An examination of boycottnfl using social network analysis. Computers in Human Behavior 115, 106623. URL: <https://www.sciencedirect.com/science/article/pii/S0747563220303708>, doi:<https://doi.org/10.1016/j.chb.2020.106623>.
- [7] Deng, J., Dong, W., Socher, R., jia Li, L., Li, K., Fei-fei, L., 2009. Imagenet: A large-scale hierarchical image database, in: In CVPR.
- [8] Hohenecker, P., Lukasiewicz, T., 2017. Deep learning for ontology reasoning. arXiv preprint arXiv:1705.10342 .
- [9] Jia, Y., Shelhamer, E., Donahue, J., Karayev, S., Long, J., Girshick, R., Guadarrama, S., Darrell, T., 2014a. Caffe: Convolutional architecture for fast feature embedding. arXiv:1408.5093.
- [10] Jia, Y., Shelhamer, E., Donahue, J., Karayev, S., Long, J., Girshick, R.B., Guadarrama, S., Darrell, T., 2014b. Caffe: Convolutional architecture for fast feature embedding. CoRR abs/1408.5093. URL: <http://arxiv.org/abs/1408.5093>.
- [11] Krizhevsky, A., Sutskever, I., Hinton, G.E., 2012a. Imagenet classification with deep convolutional neural networks, in: Pereira, F., Burges, C.J.C., Bottou, L., Weinberger, K.Q. (Eds.), Advances in Neural Information Processing Systems 25. Curran Associates, Inc., pp. 1097–1105. URL: <http://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf>.
- [12] Krizhevsky, A., Sutskever, I., Hinton, G.E., 2012b. Imagenet classification with deep convolutional neural networks, in: Pereira, F., Burges, C.J.C., Bottou, L., Weinberger, K.Q. (Eds.), Advances in Neural Information Processing Systems 25. Curran Associates, Inc., pp. 1097–1105.
- [13] Lazzez, O., Ouarda, W., Alimi, A.M., 2016. Understand me if you can! global soft biometrics recognition from social visual data, in: Proceedings of the 16th International Conference on Hybrid Intelligent Systems (HIS 2016), Marrakech, Morocco, November 21–23, 2016., pp. 527–538.
- [14] Li, L.-J., Fei-Fei, 2007. What, where and who? classifying event by scene and object recognition .
- [15] Li, J., Wu, N., Feng, Z., 2018. Model-based non-gaussian interest topic distribution for user retweeting in social networks. Neurocomputing 278, 87 – 98. Recent Advances in Machine Learning for Non-Gaussian Data Processing.
- [16] Liu, Z., Luo, P., Qiu, S., Wang, X., Tang, X., 2016. Deepfashion: Powering robust clothes recognition and retrieval with rich annotations .
- [17] Musto, C., de Gemmis, M., Semeraro, G., Lops, P., 2017. A multi-criteria recommender system exploiting aspect-based sentiment analysis of users' reviews, in: Proceedings of the Eleventh ACM Conference on Recommender Systems, Association for Computing Machinery, New York, NY, USA. p. 321–325. URL: <https://doi.org/10.1145/3109859.3109905>, doi:10.1145/3109859.3109905.
- [18] Rich, J., Haddadi, H., Hospedales, T.M., 2016. Towards bottom-up analysis of social food. CoRR abs/1603.04497.
- [19] Segalin, C., Cheng, D.S., Cristani, M., 2017. Social profiling through image understanding: Personality inference using convolutional neu-

Table 7The *DeepVisInterest* framework precision measure

Class	Number of images				
	5	10	50	75	100
Activities	1	0.4	0.4	0	0
Business	1	1	1	1	0
Drink	0.6	0.8	1	1	0.08
Education	1	1	1	1	0
Entertainment	1	1	0.5	0.5	0
Events	1	1	1	1	0.1
Family	1	1	1	1	0
Fashion	0.5	0.6	0.7	0.3	0
Fitness	1	1	0	1	0
Food	0	0	0.2	0	0
Industry	0.1	0.3	0.5	0.1	0
News	0.3	0.2	0.6	0.1	0
Outdoors	0.6	0.9	1	0.3	0
People	1	1	1	1	0
Places	1	1	0.9	0.5	0
Shopping	0.6	0.9	0.5	0.8	0
Sport	1	1	0.3	0.4	0
Technology	0.3	0.2	0	0.2	0
Travel	1	1	0	0	0
Culture	0	0	1	0	0
Hobbies	1	1	0	0	0
Lifestyle	0	0	1	0	0
Relationship	1	0	1	1	0
Wellness	0	0	0	1	0

ral networks. *Computer Vision and Image Understanding* 156, 34 – 50.

- [20] Wu, B., Jia, J., Yang, Y., Zhao, P., Tang, J., 2015. Understanding the emotions behind social images: Inferring with user demographics, in: 2015 IEEE International Conference on Multimedia and Expo (ICME), pp. 1–6.
- [21] Yang, C., Wang, Y., 2019. Online social network image classification and application based on deep learning, pp. 5–10.
- [22] Yang, C., Yu, X., Liu, Y., Nie, Y., Wang, Y., 2016. Collaborative filtering with weighted opinion aspects. *Neurocomputing* 210, 185–196. URL: <https://www.sciencedirect.com/science/article/pii/S0925231216305926>, doi:<https://doi.org/10.1016/j.neucom.2015.12.136>. sI:Behavior Analysis In SN.
- [23] Yang, L., Hsieh, C., Estrin, D., 2017. Beyond classification: Latent user interests profiling from visual contents analysis. *CoRR* abs/1512.06785.
- [24] You, Q., Bhatia, S., Luo, J., 2017. A picture tells a thousand words - about you! user interest profiling from user generated visual content. *CoRR* abs/1504.04558.

Table 8
The *DeepVisInterest* framework recall measure

Class	Number of images				
	5	10	50	75	100
Activities	0.3	0.3	0.3	0.4	0.4
Business	0.7	0.7	0.7	0.6	0.4
Drink	0.6	1	0.9	0.7	0
Education	1	0.7	1	0.8	0
Entertainment	0.9	0.7	1	0.4	0
Events	1	0.6	0.6	0.7	0
Family	0.4	1	1	0.7	0
Fashion	1	0.8	1	0.6	0
Fitness	0.3	1	0.9	0.7	0
Food	0.7	0.6	0.9	0.2	0
Industry	0.4	0.8	0	0	0
News	0.7	1	0	0.6	0
Outdoors	0.5	0	0	0.5	0
People	0.5	1	1	0.2	0
Places	0.6	0.6	0.7	0.3	0
Shopping	1	1	0.4	0.3	0
Sport	0.6	0.5	0.6	0.5	0
Technology	0.5	1	0.5	0.2	0
Travel	0.1	1	0.6	0.8	0
Culture	0	0	1	0	0
Hobbies	0	0.2	0.7	0	0
Lifestyle	0	0.4	0.4	0	0
Relationship	0	0	0	0	0
Wellness	0	0	0	1	0