FOODINTEGRITY

Ensuring the Integrity of the European food chain

613688: Collaborative Project

Seventh Framework Programme

KBBE.2013.2.4-01: Assuring quality and authenticity in the food chain

Deliverable: 21.2

Title: F.I.S.HUB Validation

Author(s): IZSPLVA, POLITO, USAL

Beneficiary(s): IZSPLVA, POLITO, USAL

Date of preparation: 20/01/2018

Period covered: months 47_50

Status: version 1

Dissemination level		
PU	Public	
PP	Restricted to other participants	
RE	Restricted to a group specified by the consortium	
CO	Confidential, only members of the consortium	X

The project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement No. 613688.

Deliverable 21.2 F.I.S.HUB Validation

TABLE OF CONTENTS

1. Executive Summary

In this task we describe the validation of the F.I.S.HUB server classifier. We collected 69 pictures of species available on the market. All personnel that took pictures followed the protocol described in Deliverable 21.1 (that will also be the same one suggested to the users of the app). With the 69 photos we obtained 1725 classification experiments by labeling each photo with the available 25 labels (corresponding to the 25 species in the database). Knowing, thanks to the DNA Barcoding the correct label for each picture, we then asked the classifier to validate each pair. The validation results showed a global Accuracy of 98,9% (Sensitivity 86,15% , Specificity 99,46%). The analysis of the performances for each individual class, suggests some refinements to both the database and the classification pipeline that will be implemented in the last months of the project.

2. Description of Deliverable

Introduction

This deliverable is linked with the WP15 and describes the validation of the F.I.S.HUB classifier. According to the protocol of Figure 1, (i) fish specimen were bought in different markets, then (ii) a picture was taken for each specimen, and finally (iii) a tissue sample for each specimen was sent to a laboratory for DNA analysis.

Figure 1: Validation Photo guidelines

Test users were recruited in order to create a heterogeneous group including inexperienced users working in an Italian food company, veterinary officers or other kind of controllers (who could have different degrees of experience with fish identification), as well as common end-users at selling points (both from UK and Italy).

Specimen collection

It was not possible to find in the market all the fish species included in the validation protocol since some of them are not available in the period in which this study has been carried out. We here present the list of the 69 specimen collected:

ORDER	FAMILY	GENUS	SPECIES	n° Pictures
Clupeiformes	Clupeidae	Clupea	harengus	5
Clupeiformes	Clupeidae	Sardina	pilchardus	7
Clupeiformes	Engraulidae	Engraulis	encrasicolus	6
Gadiformes	Merluccidae	Merluccius	merluccius	7
Perciformes	Sparidae	Dentex	dentex	3
Perciformes	Sparidae	Dentex	gibbosus	3
Perciformes	Sparidae	Pagellus	bogaraveo	1
Perciformes	Sparidae	Pagellus	erythrinus	6
Perciformes	Sparidae	Pagrus	caeruleostictus	1
Perciformes	Sparidae	Pagrus	pagrus	5
Pleuronectiformes	Pleuronectidae	Microstomus	kitt	5
Pleuronectiformes	Pleuronectidae	Pleuronectes	platessa	5
Pleuronectiformes	Scophthalmidae	Psetta	maxima	8
Pleuronectiformes	Soleidae	Solea	vulgaris	7

DNA Barcoding

In IZSPLVA and USAL the DNA was extracted from 25mg of tissue using a commercial kit based on silica purification (ReliaPrep™ gDNA Tissue Miniprep System, Promega). All samples were analysed amplifying a portion of COI gene in a PCR reaction containing Platinum™ Quantitative PCR SuperMix-UDG 1X (Invitrogen) and 0.3 µM of each primer taken from Ward et al. (2005), but assembling the two primers forward and the two reverse into a set of degenerated primers. Amplicons were sequenced using Sanger method on both strands and consensus sequences were compared to those reported in public databases BLAST (http://blast.ncbi.nlm.nih.org) and in BOLD Identification System (www.boldsystems.org) too. The species was assigned on the basis of a similarity with sequences reported in public databases ≥ 98%.

FISHUB Classification pipeline

The prototype of the F.I.S.HUB mobile application was supplied to people involved in the validation. The pictures were taken using the app running on mobile devices; the picture analysis phase was done off-line to get a more detailed result than the one that will be shown by the app (even if the classification algorithm is exactly the same). The classifier used in the validation process has been described in the Deliverable 15.2. Each picture is annotated with a label and the classifier role is to validate the label against the picture. The classifier is able to provide a score between 0% and 100% for each labeled photo. In this way, for each picture we collected 25 scores. We considered "50%" as the experimental cutoff value to assign the consensus or to highlight a mislabeling.

3. **Results**

To compute the overall classifier performances we asked the FISHUB server to validate each photo against all 25 labels. This allowed to perform **1725 classification experiments**.

The validation performance of the classifier is as follows:

VALIDATION	FISHUB confirms LABEL	FISHUB predicts fraud
The fish is correctly LABELLED	55	10
The fish is MISLABELED	10	1650

Measure	Value [%]	Derivations
Sensitivity	86,15	TPR = TP / (TP + FN)
Specificity	99,46	SPC = TN / (FP + TN)
Precision	86,15	PPV = TP / (TP + FP)
Negative Predictive Value	99,46	NPV = TN / (TN + FN)
False Positive Rate	0,54	FPR = FP / (FP + TN)
False Discovery Rate	13,85	FDR = FP / (FP + TP)
False Negative Rate	13,85	FNR = FN / (FN + TP)
Accuracy	98,96	ACC = (TP + TN) / (P + N)
F1 Score	86,15	F1 = 2TP / (2TP + FP + FN)
Matthews Corr. Coeffi.	85,61	TP*TN - FP*FN / sqrt((TP+FP)*(TP+FN)*(TN+FP)*(TN+FN))

The detailed results of the validation are summarized in the table below. The meaning of each column is as follows:

- Order/Family/Genus/Species: the declared label in the marketplace
- n° Experiments: number of pictures taken for that label multiplied by the number of possible labels (i.e. 25).
- TP: the number of correctly labeled pictures correctly validated.
- TN: the number of mislabeled pictures correctly validated.
- FP: the number of correctly labeled pictures that were not correctly validated.
- FN: the number of mislabeled pictures that were not correctly validated.
- Sens: specificity of the classifier for that species
- Spec: sensitivity of the classifier for that species

ORDER	FAMILY	GENUS	SPECIES	n° Experiments	TP	TN	FP	FN	Sens.	Spec.
Clupeiformes	Clupeidae	Clupea	harengus	125	0	123	1	1	0,00%	99,19%
Clupeiformes	Clupeidae	Sardina	pilchardus	175	6	167	1	1	85,71%	99,40%
Clupeiformes	Engraulidae	Engraulis	encrasicolus	150	6	144	0	0	100,00%	100,00%
Gadiformes	Merluccidae	Merluccius	merluccius	175	6	167	1	1	85,71%	99,40%
Perciformes	Sparidae	Dentex	dentex	75	2	71	1	1	66,67%	98,61%
Perciformes	Sparidae	Dentex	gibbosus	75	1	70	2	2	33,33%	97,22%
Perciformes	Sparidae	Pagellus	bogaraveo	25	1	24	0	0	100,00%	100,00%
Perciformes	Sparidae	Pagellus	erythrinus	150	6	144	0	0	100,00%	100,00%
Perciformes	Sparidae	Pagrus	caeruleostictus	25	0	23	1	1	0,00%	95,83%
Perciformes	Sparidae	Pagrus	pagrus	125	3	118	2	2	60,00%	98,33%
Pleuronectiformes	Pleuronectidae	Microstomus	kitt	125	4	119	1	1	80,00%	99,17%
Pleuronectiformes	Pleuronectidae	Pleuronectes	platessa	125	5	120	0	0	100,00%	100,00%
Pleuronectiformes	Scophthalmidae	Psetta	maxima	200	8	192	0	0	100,00%	100,00%
Pleuronectiformes	Soleidae	Solea	vulgaris	175	7	168	0	0	100,00%	100,00%

The DNA Barcode analysis detected 9 mislabeled fishes:

- n° 5 Clupea harengus barcoded as Sardina pilchardus
- n° 1 Dentex dentex barcoded as Dentex gibbosus
- n° 2 Dentex Gibbosus barcoded as Pagrus major
- n° 1 Pagrus pagrus barcoded as Pagrus major

The F.I.S.HUB classifier labeled as frauds 10 correctly labeled fishes (FN):

- n° 1 Clupea harengus that F.I.S.HUB classified as Sardina pilchardus
- n° 2 Sardina pilchardus that F.I.S.HUB classified as Clupea harengus
- n° 1 Merluccius merluccius that F.I.S.HUB classified as Merlangius merlangus
- n° 1 Dentex dentex that F.I.S.HUB classified as Dentex gibbosus
- n° 2 Dentex gibbosus that F.I.S.HUB classified as Dentex dentex
- n° 1 Pagrus caeruleostictus that F.I.S.HUB classified as Pagrus pagrus
- n° 1 Pagrus major that F.I.S.HUB classified as Pagrus pagrus
- n° 1 Microstomus kitt that F.I.S.HUB classified as Limanda limanda

From the available results there is no clue of any correlation between the classifier performances and variables like the person who took the pictures or the location where the pictures were taken. One parameter that showed a strong correlation with performances was the picture background. All pictures taken not following the protocol (concerning the white and luminous background) were misclassified. These pictures were not taken into account in the validation results and have been removed from the validation dataset.

4. Discussion and conclusions

The DNA Barcoding revealed that 9 out of 69 (~13%) of the collected specimens were mislabelled (all the five Clupea harengus, one Dentex dentex, two Dentex gibbosus and a Pagrus pagrus). This result requires to be carefully analyzed because it might reveal a problem in the labeling at the source; it is probably not a fraud but just a common mistake in labeling species that are highly similar. Because of this mislabeling issue, it is also possible that some of the pictures of the database are mislabeled and for this reason a reliable classification is more difficult. Fortunately this problem seems to be affecting only two or three classes. In the final month of the project we will investigate this issue and possibly rearrange the classes in the database to obtain more reliable classifications.

From the classification scores obtained in the validation two main issues arose:

1) some of the classes in the database appear not to be large enough to allow a reliable classification using the Deep Learning approach. The majority of mislabelling not detected by the F.I.S.HUB classifier is for Dentex and Pagrus genus. To improve the quality of the database, and therefore the classifier performances, we are designing the server-side of the App so that the correctly classified pictures will be added (after a manual check by professionals) to the database, and the classifier will be periodically re-trained with the new pictures. This will allow to greatly improving the performances of the classifier over time.

2) the classification pipeline is not performing optimally for some species which are in fact very similar (i.e. Dentex dentex, Dentex gibbosus, Pagrus major, Pagrus pagrus).

To address this problem we are training one-class siamese classifiers so that they will be able to better recognize the small differences which are, right now, the cause of most misclassifications.