JRC TECHNICAL REPORT

Use of geotagged photographs to evidence Land Cover and Land Use across EU policies

Loudjani, P., Sima, A., and Devos, W.

2023

Contact information
Name: Philippe LOUDJANI
Address: European Commission
Joint Research Centre
Directorate D – Sustainable Resources
Food Security Unit
TP272, 27B 00/32
Via Enrico Fermi 2749, I-21027 Ispra / Italy
Email: Philippe.loudjani@ec.europa.eu
Tel.: +39 0332-786160

EU Science Hub
https://joint-research-centre.ec.europa.eu

JRC135007

Ispra: European Commission, 2023

How to cite this report: Loudjani, P., Sima, A., and Devos, W. *Use of geotagged photographs to evidence Land Cover and Land Use across EU policies*, Ispra, 2023, JRC135007

Contents

Abstract

As part of supporting work of anticipation and implementation of EU policies, JRC operates surveys of potential innovative tools and new technologies. In recent years, the use of geotagged imagery appeared to be a tool offering a wide range of potential applications.

In a first technical report (JRC 120223), the Common Agricultural Policy (CAP) has been the main policy targeted since the European Commission adopted new rules allowing for a range of modern technologies to be used when carrying out checks for CAP payments. Following this, there has been a significant uptake of the use of geotagged by Member States (MSs) paying agencies and significant developments through studies or projects confirming the usefulness of geotagged photos for the CAP support on land related elements. It also allowed identifying potential for non-land related practices/elements, including insurance schemes or animal related requirements and for support to other policies that the CAP e.g. Climate and Environment.

Thanks to the ease of use and flexibility of smartphones and digital cameras, the range of practices, land uses and land covers that geotagged photos can depict through a multitude of sceneries, landscape elements or single plants/plant elements is almost unlimited.

Nevertheless, the actual handling and management of a wide range of situations and above all a high number of photos may rapidly become very challenging and cumbersome. It is thus of high importance to dedicate time in a preparatory phase to identify the use cases where geotagged photos will offer a clear advantage compared to other solutions. Attention should be paid also to solutions to facilitate and/or automatise the necessary steps of photo handling and management.

This report intends to provide the necessary technical guidance and best practices to efficiently deploy geotagged photos of land cover / land use (but not only) elements in the frame of different EU policies requirements.

Elements provided in the different chapters of this report provide knowledge and best practices to help organising the sound acquisition and use of geotagged photos through the creation and access to photo collections, the creation and training of classification and/or segmentation algorithms, the development of apps embedding user-friendly and intuitive solutions to guide operators from data capture up to data delivery to the administration data servers.

Foreword

This report intends to provide practical recommendations and some necessary best practices to efficiently use geotagged photos to evidence a wide range of land cover / land use elements as part of EU policy requirements.

Acknowledgements

The authors wish to acknowledge the experts from paying agencies of Catalunia, Spain (especially Laura PAVON), Belgium-Flanders (especially Peter ROGGERMANS), and Bulgaria (especially Rositsa IVANOVA) for providing data and input.

Authors

LOUDJANI Philippe, SIMA Aleksandra, and DEVOS Wim.

Executive summary

This report intends to provide the necessary technical guidance and best practices to efficiently deploy geotagged photos of land cover / land use (but not only) elements in the frame of different EU policies requirements.

As part of supporting work of anticipation and implementation of EU policies, JRC operates surveys of potential innovative tools and new technologies. In recent years, the use of geotagged imagery appeared to be a tool offering a wide range of potential applications.

In a first technical report (JRC 120223), the Common Agricultural Policy (CAP) has been the main policy targeted since the European Commission adopted new rules allowing for a range of modern technologies to be used when carrying out checks for CAP payments. It included the possibility to use geotagged photos. Following this, there has been a significant uptake of the use of geotagged by Member States (MSs) paying agencies and significant developments through studies or projects confirming the usefulness of geotagged photos for the CAP support on land related elements. It also allowed identifying potential for non-land related practices/elements, including insurance schemes or animal related requirements and for support to other policies that the CAP e.g. Climate and Environment.

Thanks to the ease of use and flexibility of smartphones and digital cameras, the range of practices, land uses and land covers that geotagged photos can depict through a multitude of sceneries, landscape elements or single plants/plant elements is almost unlimited.

Nevertheless, the actual handling and management of a wide range of situations and above all a high number of photos may rapidly become very challenging and cumbersome. It is thus of high importance to dedicate time in a preparatory phase to identify the use cases where geotagged photos will offer a clear advantage compared to other solutions (e.g. use of Sentinel temporal series), to clearly identify what should be evidenced and define scenarios and workflows the most adapted to the final users. Attention should be paid also to solutions to facilitate and/or automatise the necessary steps of photo handling and management even if, in some situations, the choice of a simple pragmatic solution can be more efficient (in terms of time and cost).

Elements provided in the different chapters of this report provide knowledge and best practices to help organising the sound acquisition and use of geotagged photos through the creation and access to photo collections, the creation and training of classification and/or segmentation algorithms, the development of apps embedding user-friendly and intuitive solutions to guide operators from data capture up to data delivery to the administration data servers.

Regardless of the chosen solution, the key to the success is getting the stakeholders, beneficiaries and photo providers actively involved in the process and collaborating. The developed apps and data processing solutions should be endorsed and used by the targeted community. The need for this involvement can be further increased through awareness-raising campaigns (using a variety of communication channels such as emails, newsletters, social media …) and by ensuring training sessions.

The sharing of experience and good practices is an asset to facilitate the use of geotagged photos. Such sharing is the main purpose of this report while being aware that many of the underpinned technologies and tools are developing very fast and any guidance will have to be updated accordingly.

1 Introduction

As part of supporting work of anticipation and implementation of EU policies, JRC operates surveys of potential innovative tools and new technologies. In recent years, with the rapid development of the market and technology of digital cameras, smartphones and other devices with cameras equipped with Global Navigation System (GNSS) antenna, the use of geotagged imagery appeared to be a tool offering a wide range of potential applications.

In the first version of this technical report (JRC 120223), the Common Agricultural Policy (CAP) has been the main policy targeted since, as part of its move to simplify and modernise the EU's CAP, the European Commission adopted new rules allowing for a range of modern technologies to be used when carrying out checks for CAP payments. This included the possibility to replace physical checks on farms (On-The-Spot Checks, OTSC) with a system of automated checks based on analysis of Earth observation data. The proposed monitoring approach (Checks by Monitoring CbM) uses mainly the earth observation data provided by the Copernicus Sentinel satellites and also includes the possibility to use geotagged photos to support and complement OTSC and CbM methods when they do not lead to conclusive results. They can also be used as ground truth information provided by farmers or other stakeholders for Integrated Administration and Control System (IACS) processes such as the update of the Land Parcel Identification System (LPIS). The possible use of geotagged photos has been renewed and extended as part of the CAP2020+ legislation.

There has been already a significant uptake of the use of geotagged photos by Member States (MSs) paying agencies and significant developments through studies or projects (e.g. NIVA, DIONE …) confirming their usefulness for the CAP support on land related elements. But it also allowed identifying potential for non-land related practices/elements, including insurance schemes or animal related requirements and for support to other policies that the CAP e.g. Climate and Environment.

This report intends to provide the necessary technical guidance and best practices to efficiently deploy geotagged photos of land cover / land use (but not only) elements in the frame of different EU policies requirements.

2 Potential for geotagged photos

After their introduction in the CAP legislation in 2018, geotagged photos have started to be used in a wide range of situations in the CAP checks and management context to serve as input to:

- document Rapid Field Visits (RFV);
- provide parcel contextual information to update the Land Parcel Identification System (LPIS);
- evidence elements not monitorable with satellite imagery;
- follow-up selected inconclusively monitored parcels;
- provide ground truth for quality assessment or training of machine learning processes.

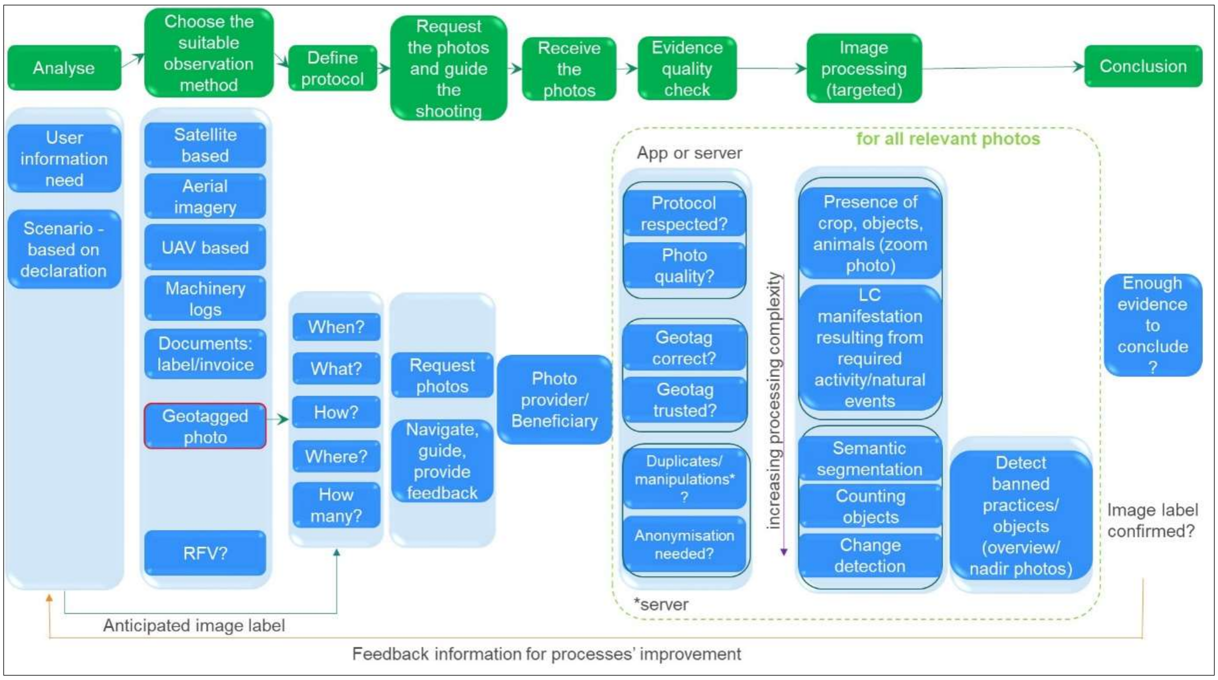
In fact, thanks to the ease of use and flexibility of smartphones and digital cameras, the range of objects, situations, actions that can be documented and captured through geotagged photos is almost unlimited. From a large rural landscape area up to a close-up photo of a wheat's head, from a meadow just mowed to a peatland area in a Natura2000 zone, from cow ear tag to a machine spreading plant protection product, vast are the support possibilities in different policies (e.g. AGRI, CLIMA, ENV …). Nevertheless, to not be rapidly and unnecessarily overwhelmed by a huge quantity of photos to process and handle, it is highly recommended to identify elements and focus on developments for which the use of geotagged photos will bring a clear advantage compared to the use of other types of technologies/tools.

3 General processing for any ground elements to be evidenced

There are many policies which require gathering information on land cover and/or land use. The CAP constitutes the main one with such requirements and this policy will often be used in this report to identify use cases.

Using experiences gathered so far, we propose hereafter a general workflow to deal with any request for evidence.

Figure 1. Proposed general workflow to handle requests for information to be evidenced on the ground.



Source: JRC

Seemingly, it sounds easy to take a picture and repeat this for an almost unlimited list of objects, features or situations. Nevertheless, the actual handling and management of a wide range of situations and above all a high number of photos may rapidly become very challenging and cumbersome. It is thus of the utmost importance to scout the available observation methods and choose the geotagged photos in use cases where this tool will bring an advantage compared to other technical solutions. Attention should be paid also to solutions to facilitate and/or automatise the necessary steps of photo handling and management.

3.1 Choice of the suitable observation / data capturing method

Prior to focusing on the use of geotagged photos, it is strongly recommended to identify and analyse the need(s). The most suitable observation method/tool should be selected based on the nature of the user information need and the character of the changes/states to be evidenced, considering the data availability and ease of access to them. For example, the identification of a 10m wide buffer strip with unharvested crop (wild animals forage) will hardly be identifiable with Sentinel data, but relatively easy to confirm with: aerial imagery (if captured in the right period), drone imagery or geotagged photo. On the contrary, it would be cumbersome to systematically require pictures of cereals fields when most of them are rather accurately identifiable automatically with Sentinel data (*not only limited to cereals, there is now a wide range of papers describing automatic classification methods of temporal series of Sentinel 2 (optical) and/or Sentinel 1 (SAR) data/indices and reporting high crop classification accuracy*).

Use cases need to be closely analysed, possibly together with the information that can be anticipated from the user declaration (the GeoSpatial Application, GSA).

In the frame of the CAP, it is advised for instance to draw the list of all use cases (interventions, practices, and land cover types) included in the national strategic plan that will condition farms' activities. Once the list is established, a first screening should be done in order to determine all the use cases that, according to current technical knowledge and data availability, could be evidenced using Sentinel temporal information or other imagery sources (single observation or with limited revisit frequency), i.e. VHR imagery or drone surveys.

In the cases when geotagged photos seem to be the optimal data source, further analyses and split should be done according to the type/nature of activity/state/phenomena that needs to be evidenced. For example, to evidence afforestation, terracing and other large-scale land/farm reorganisation, photos of the state before and after should be requested in order to facilitate understanding of the actual change. For cases linked with biodiversity, the photos should be captured in the period when the biodiversity indicators are visible, e.g. when the flower species of interest are blooming.

Based on such analyses, detailed protocols for data capturing, collection and processing should be defined, with the intention in mind to minimise user (e.g. farmer and/or administration) efforts and maximise the information gain.

Also, in order to optimise the chances to get the relevant high quality information on time, i.e. the geotagged photo captured in the right moment, from the requested position and in a requested way, guidance and assistance should be provided upfront to the user by various means (e.g. guiding text or a vocal message with clear step-by-step instructions in App, email sent to the user …).

3.2 Defining protocols

For each use case, a general description of the timeline of phenomena and activities to be evidenced should be documented. This step is analogous to the scenario definition in the Checks-by-Monitoring methodology document (JRC127678), which serves to facilitate the selection of the most relevant and reliable markers and their operating time range. In the case of the geotagged photos, the scenario aims at providing information on what to photograph, when, how to take pictures, where, and how many should be taken.

The what?

Depending on the use case, one has to define what element is essential to capture to evidence the need. It could be a specific flower to indicate some grassland richness, the result of a specific activity such as grass left on the ground, a picture evidencing the presence of plant mixture in grassland or even the presence of animals on a parcel.

The when?

The timing for the photo capturing should be defined in order to obtain unambiguous evidence of a targeted element (flowering plant, right after an activity, just before or just after a fixed deadline …).

The how?

Depending on the use case, the operator should be guided to take a ‘close up’ or ‘landscape’ photography, the viewing angle may be horizontal, oblique or nadir.

The where?

It can be requested to take a picture from a specific location (e.g. corner of a parcel, middle of a parcel, to evidence a specific element in or bordering a parcel …). In some situations, the decision of the location may be left to the operator.

The how many?

Depending on the situation, information can be provided on how many pictures should be taken from each viewpoint. For instance, some situations may require a picture of the four cardinal directions (to provide a general view of the parcel environment), other may require both nadir and landscape views (to ensure that a parcel is fully covered by a specific type of cover and/or check its density, species composition etc.).

For the last 2 elements (where, how many), the specifications will largely depend on the characteristics of the terrain/parcel and on cost/effectiveness considerations.

3.3 Main prerequisites

As already stated, taking a picture appears to be an easy action. However, getting the right picture(s) at the right location and the right time and in due time is something far from being a given. In order to foster the deployment of geotagged photos, we have identified 3 necessary preconditions i.e. having the operator pro-active in the process, providing methodological recommendations on how to capture the photo so that it contains the required evidence, and providing methods to ease and automatise steps of photo processing.

The **Figure 2** below is an attempt to summarise all components to consider in photo handling and processing together with the possible levers to ease and automatise them.

Elements evidenced in bold blue represent the minimum components that have to be in place for geotagged photo handling i.e.:

- an App to at least capture geotagged photos but that can contains much more functions such as guidance to the parcel of interest, guidance to element of interest, photo quality checking or even photo content checking;

- A geotagged photo database to at least store photos provided by users but possible also to be used to train photo content recognition algorithm;

- a procedure of photo quality assessment to ensure that photos sent by users accurately evidence the elements requested.

Elements evidenced in red bold text represent the minimum information that have to be recorded for each geotagged photo i.e.:

- the date, time, and location of photo captured;

- The identification of the operator who took the photo;

- tags/labels for the considered photo.

Elements framed with a blue dashed line evidence components or actions intending to improve photo and photo content quality and fit-for-purpose such as:

- the provision of guidance to take the right picture;

- the use of methods to improve the positioning accuracy and the photo integrity;

- the setting of a photo Quality Assessment process;

- the building of a curated photo Database to train and develop automatic photo processing algorithms.

Elements framed with a green dotted line evidence components or actions intending to ease the fit-for-purpose photo obtaining such as:

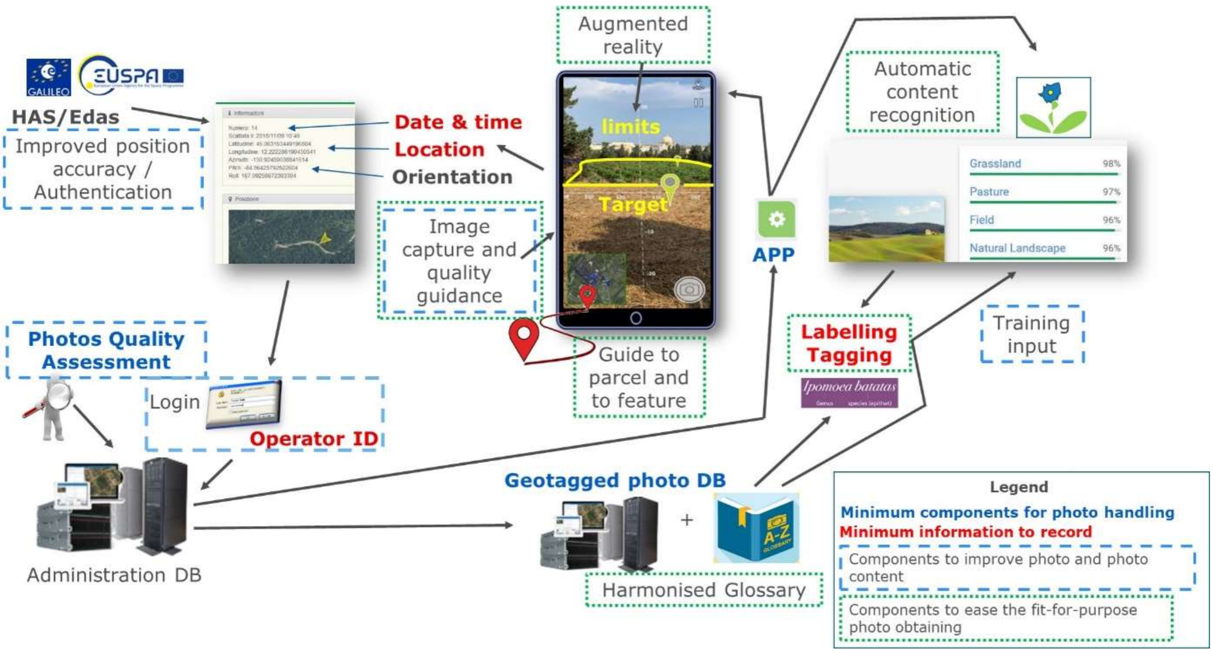
- guiding the operator to the right point of interest;

- guiding the operator to capture the right element;

- use automatic photo content recognition and/or labelling/tagging.

More details on all these elements are provided later in this report.

Figure 2. Summary of all components to consider in photo handling and processing together with the possible levers to ease and automatise them.



Source: JRC.

3.3.1 Pro-active role of operators (e.g. farmers)

As previously mentioned, the use of geotagged photos has potential in the frame of several EU policies. Nevertheless, to date, the CAP is still the policy with the main explicated interest and consequently, farmers or farmers' representatives (i.e. farmers associations, farm advisors etc.) constitute the main photo capturers and providers.

When farmers are requested to capture photos evidencing the presence of an object or activity conditioning the eligibility for a specific subsidy, it is crucial for the processing of the request that they act promptly and send the evidencing photo(s) in due time. If not delivered in time, the administration could be forced, as a last resort, to perform a field visit with a high risk of being too late in the season to still find the relevant evidence.

It is thus of the utmost importance that farmers endorse these processes and play a pro-active role in providing photos. Nevertheless, it is currently a fact that in many Member States the involvement of farmers in the IACS process is still very limited. Not only for geotagged photos, but for the farm dossier processing as a whole, the active role of farmers is needed (acting, reacting to warning messages or request to modify aid application etc.). That means that a change in habits and culture has to be operated. It involves communicating with farmers in a manner that encourages their endorsement of processes through for instance training or learning sessions. It also involves setting easy to use an information exchange system with farmers while ensuring the safety and security of information transfer.

Once the concept and usefulness accepted by farmers/operators, it is possible to think that they may anticipate situations that might be challenging and capture photos when considered relevant, but would only upload selected photos to the server at a later date, upon request.

The minimum requirements and functionalities of tools and apps for geotagged photos are addressed in the following chapters. This information is intended to be used by administrations to produce and provide farmers with a guideline/training on best practices for data collection in a suitable form (video, in app tutorial, short document).

3.3.2 Virtually enhanced instructions to capture photos fit-for-purpose

It is of high interest to provide sound instructions to the operator/farmer so that administrations will receive and process only useful and fit-for-purpose photos. Processing and rejecting unworkable photos may rapidly generate costs (in time and energy) and frustration on the side of the photo providers (i.e. farmers). In some cases, it may also not be possible to re-take a picture since the element to evidence may not be any more present in the field.

Levers to capture the fit-for-purpose photo are manifold. They can consist of:

- Navigation assistance

Assistance can be provided to get to the parcel of interest and/or provision through augmented reality with the provision of a pin or limits to indicate the location and direction of the photo to be taken.

- Guidance for image content and framing

It can consist of a text or voice message (possibly complemented by a photo example) describing the element(s) to be captured (crop, field, practice, phenological stage …), which type of photo to capture (close up, landscape …) and how many.

More details are provided later in this report.

3.3.3 Automation of image capturing and processing

With increased acceptance of the new way of documentation of agricultural practices and thus broader participation of the photo providers (i.e. farmers), especially closer to the deadlines for evidencing the declared activity, the quantity of imagery to be processed by an administration/entity may rapidly be so high that it would prevent their exhaustive visual check. Thus, technical solutions are necessary to process and validate photos in an automated way at different steps of image capturing and processing.

3.3.3.1 Ensuring image quality

Automatic warning and/or blocking of image capturing may be implemented if backlight is detected or in case of over/under light exposure or even if the image is blurred (e.g. motion blur).

3.3.3.2 Ensuring image integrity

Techniques exist to detect if photos have been possibly manipulated (date, location, image content) or duplicated (photo of another photo, non-original photo). Efforts in these techniques are now boosted by the need to detect 'false-realistic' images (e.g. Hugging face, Tanaka et al., 2021; Ankolekar et al., 2022,) generated by newly available AI systems from a description in natural language like DALL-E 2 or Midjourney.

One can note that some administrations have developed dedicated smartphone/tablet applets using combinations of solutions to avert tampering with data. As example, a photo is blocked for upload if not sent within few seconds after capture, thus not leaving time for any image manipulation (and in case a GSM network is available for transfer at the location of picture capture).

3.3.3.3 Automation of image content recognition

Automatic screening of photos is essential. With the recent advances in computer vision and machine learning techniques, image content recognition has become a realistic solution, not only to confirm/contradict alleged crop types/activities depicted in the photo but also to filter out unwanted objects in the photographs.

3.3.3.4 Automated photo labelling and tagging

To date, most of the algorithms for automatic recognition and classification of image content rely on large training image datasets. The quality of the training dataset highly depends on the correctness and quality of image tagging (and/or labelling). The training datasets are often tagged and labelled manually based on visual interpretation. This step is very tedious and time-consuming. Thus, ensuring precise and correct assignment of image tags at photo capture or at very last at photo storage in the database is of great importance.

4 Technical details

After having set the scene for the possible use of geotagged photos to evidence a wide range of land cover / land use elements according to the EU policy requirements, this chapter intends to provide corresponding complementary technical details and best practices.

4.1 Geotagging

Geotagged photos are commonly referred to as digital photographs with embedded spatial information. Geotagging of the photos can be done manually or automatically. In short, geotagging consists in saving at least the latitude and longitude coordinates into the Exchangeable Image File (Exif) data of several JPEG (Joint Photographic Experts Group) and TIFF (Tag(ged) Image File Format) file formats. Most of the new digital cameras, including the ones of smartphones, use the Exif annotation to store additional photo relevant information such as shutter speed, F number, ISO number, image resolution and compression level, whether a flash was triggered, ISO number, auxiliary lenses, etc.

With the recent fast pace of developments, most smartphones and cameras come with a built-in GNSS (Global Navigation and Satellite System) receiver that enables automatic geotagging and time-stamping based on time and position retrieved by the inbuilt antenna.

Figure 3. Example of a geotagged photo of a potato field captured in the frame of the LUCAS survey 2018, together with the detailed metadata registered in its Exif. Note that most of the information is automatically derived from device and GNSS antenna data. Image description, artist and copyright information have been added 'manually'.

Source: European Union, LUCAS survey2018.

EXIF

Make	Apple
Model	iPhone 6
Orientation	Horizontal (normal)
ResolutionUnit	inches
Software	11.3.1
ModifyDate	2018:06:11 09:53:27
ImageDescription	LUCAS 2018, 35303726, Potatoes for LC1, Not relevant for LC2, null for LU1, Not relevant for LU2
Artist	UKSU005
Copyright	(c) European Union, 2015 - Reuse authorised - The reuse policy of European Commission documents is regulated by Decision 2011/833/EU (OJ L 330, 14.12.2011, p. 39) - The reuser has to acknowledge the source of the documents; has the obligation not to distort the original meaning or message of the documents; guarantee the non-liability of the Commission for any consequence stemming from the reuse.
ExposureTime	1/169
FNumber	2.2
ExposureProgram	Program AE
ISO	32

DateTimeOriginal	2018:06:11 09:53:27
CreateDate	2018:06:11 09:53:27
ShutterSpeedValue	1/169
ApertureValue	2.2
BrightnessValue	6.897630332
ExposureCompensation	0
MeteringMode	Multi-segment
Flash	Auto, Did not fire
FocalLength	4.2 mm
SubjectArea	1631 1223 1795 1077
SubSecTimeOriginal	895
SubSecTimeDigitized	895
XPTitle	[illegible]
XPSubject	[illegible]
ColorSpace	sRGB
ExifImageWidth	1600
ExifImageHeight	1200
SensingMethod	One-chip color area
SceneType	Directly photographed
ExposureMode	Auto
WhiteBalance	Auto
FocalLengthIn35mmFormat	29 mm
LensInfo	4.15mm f/2.2
LensMake	Apple

LensModel	iPhone 6 back camera 4.15mm f/2.2
GPSLatitudeRef	North
GPSLatitude	56.004222
GPSLongitudeRef	West
GPSLongitude	2.748719
GPSAltitude	26.94140127 m
GPSTimeStamp	08:53:28
GPSSpeedRef	km/h
GPSSpeed	0
GPSTrackRef	True North
GPSTrack	67.8515625
GPSDateStamp	2018:06:11
GPSHPositioningError	5 m

Source: European Union, LUCAS survey2018.

In order to maximise the usefulness of geotagged photos, at minimum the following metadata should be recorded together with a photo:

– time and date of the photo capture, preferably obtained directly from the GNSS antenna (GPSTimeStamp and GPSDateStamp in Exif);

– the geographical location of the camera at the time of photo capture, also preferably obtained directly from the GNSS antenna (GPSLatitude, GPSLongitude, GPSAltitude);

– orientation (heading) of the camera at the time of photo capture, if available (GPSImgDirection and GPSImgDirectionRef);

– the identification of the operator that can be realised by personalised access to the app (login);

– basic information on the mobile device and inbuilt camera, such as mobile device brand and model number. Such information can help to retrieve e.g. the original image dimensions or focal length of the photo, or to assess the quality of provided measures of camera position and orientation;

– the accuracy and effectiveness of the image content recognition algorithm depend on the quality of the image training set and especially on the correctness and precision of tagging (or labelling). So (automatically) storing tags and labels at photo creation is important. For this, attention should also be paid to the semantics used.

Analysing and storing other data often recorded in the Exif metadata can be of benefit as well, i.e. data indicating positioning precision such as: the elevation and the Dilution of precision (GPSDOP), horizontal positioning error (GPSHPositioningError) or information on whether the differential correction was used (GPSDifferential). Although most of the above listed metadata can nowadays be automatically recorded by modern photo cameras, a dedicated app for mobile devices should be developed to ensure information integrity and security of the provided data.

4.2 Position accuracy

Geolocation is the estimation of the real-world geographic location of an object. In smartphones, geolocation can be based on information about nearby telecommunication towers and WiFi nodes or, more and more frequently, directly from GNSS signals.

The accuracy of a location derived from the telephone service provider’s network infrastructure depends on the density of base stations (“antennas”) in the area and the development advancement of the timing solutions used by the network. This geolocation technique is based on network triangulation so that the location of the device is estimated based on estimated distances to the nearest base stations. The accuracy of such solutions, relying solely on the cellular network, ranges from one to several km (Seyyedhasani, Dvorak, Sama, & Stombaugh, 2016).

The accuracy of the GNSS positioning is influenced by other factors. The satellites broadcast their signals from known positions in space with a certain accuracy, but the signal reaching the receiver is influenced by multiple factors including satellite constellation geometry, signal blockage and multipath reflectance or atmospheric conditions. The quality of the receiver's antenna and the algorithms processing the captured signal are of high importance as well.

Despite the fact that geolocation in smartphones is based on GNSS observations combined with Wi-Fi and cellular tower information, as well as an online database of satellite locations (A-GPS), the positioning accuracy of a smartphone is still considerably worse than that of dedicated GNSS devices designed solely for survey or navigation purposes. This is due to the fact that most smartphone GNSS antennas use linear polarization (Zhang, Tao, Zhu, Shi, & Wang, 2018), making it much more prone to multipath effects from GNSS signals reflected by nearby surfaces, than it is in the case of more advanced GNSS receivers (designed to minimize the multipath effect). Performance of the smartphones' GNSS antennas may also be jeopardised by their suboptimal placement inside the device (Roberts et al., 2018). A typical 'smart' device usually includes antennas for GNSS, Wi-Fi, Bluetooth, and different motion, environmental and position sensors that are expected to operate simultaneously without interfering with one another, and the GNSS sensor is not necessarily prioritised in solving such device design puzzle (in opposite to the devices designed solely for navigation or positioning). On top of that, to save power, the GNSS receiver can be turned off between readings (duty cycle) further deteriorating the accuracy of positioning (Roberts et al., 2018).

Based on the above mentioned points, it becomes inevitable that the positioning performance of the smartphone varies significantly between brands and models. Nevertheless, ranges of achievable positioning accuracies are reported in the literature. A non-exhaustive list of studies exploring the potentials of smartphones for point positioning and the resulting 2-Dimensional Root Mean Square Errors (2DRMSE) at 95% probability) is presented in **Table 1.** A non-exhaustive list of positioning accuracies obtained using smartphone devices. Errors reported originally by the authors were in several cases recomputed to derive 2DRMSE reported in this table.

Table 1. A non-exhaustive list of positioning accuracies obtained using smartphone devices. Errors reported originally by the authors were in several cases recomputed to derive 2DRMSE reported in this table.

Type of observation	2DRMSE [m]	Environment	Source
Single frequency GPS	<7	Open horizon	(Robustelli, Baiocchi, & Pugliano, 2019)
Single frequency GPS	<8	Open horizon	(Seyyedhasani et al., 2016)
Single frequency GPS	<10	Open horizon, marine environment	(Specht, Dabrowski, Pawelski, Specht, & Szot, 2019)
Single frequency GPS	<13	Urban canyon	(Robustelli, Baiocchi, & Pugliano, 2019)
Single frequency Galileo	<14 E1 <5 E5a	Open horizon	(Robustelli et al., 2019)
Single frequency GPS+ Glonass	< 12 (except one device: 32)	Open horizon	(Musulin, Kos, & Brčić, 2014)
Single frequency GPS+ Glonass	<7	Open horizon, marine environment	(Specht et al., 2019)
Single frequency GPS+ Glonass + Beidou	<12	Open horizon, marine environment	(Specht et al., 2019)
Single frequency GPS+ Glonass	<5	Open horizon	(Tomaštík, Tomaštík, Saloň, & Piroh, 2016)
Single frequency GPS+ Glonass	<34	Inside forest, leaf-on	(Tomaštík et al., 2016)
Single frequency GPS+ Glonass	<16	Inside forest, leaf-off	(Tomaštík et al., 2016)
Single frequency GPS+ Glonass + Galileo	<5	Open horizon	(Robustelli, Baiocchi, & Pugliano, 2019)

Single frequency GPS+ Glonass + Galileo	<12	Urban canyon	(Robustelli et al., 2019)
dual frequency Galileo (E1,E5) + GPS (with external antenna)	< 2.8	Open horizon	(Roberts et al., 2018)
Single frequency GPS+ Glonass + Galileo + Beidou	<2.4	Open horizon	(Dabove et al., 2020)
Single frequency GPS	<10	Open horizon	(Dabove et al., 2020)

Source: JRC

In Android 7+ operating devices, developers have been given more flexible access to the raw GNSS data received by the chipset. More complex data filtering and computation techniques, that have been used for years in surveying equipment, are now being deployed in smartphone applications to improve positioning accuracy. One example of such an algorithm was presented and implemented in the eGNSS4CAP project (GSA, 2019a).

In January 2023, the EU's Galileo GNSS started delivering its High Accuracy Service (HAS) that improves the accuracy of Galileo to sub-meter levels, with a typical positioning accuracy (95%) of about 20cm horizontally and 40cm vertically. For smartphones, the corrections provided by the service will be available at a later stage via internet streaming (HAS is distributed via E6b Galileo frequency, unavailable in the GNSS smartphone mounted chipsets).

The Open Service Navigation Message Authentication (OS-NMA) service is in the final stage of public testing (signal in space available). Nevertheless, despite the European law enforcing that all smartphone devices distributed in the EU are Galileo-enabled, the complexity of OS-NMA processing on the receiver side requires significant efforts and time before this new technology will be fully embraced by the mass market smartphone producers.

Regarding smartphones equipped with dual-frequency GNSS receivers, despite the very promising results (sub-meter accuracy) obtained with similar chipsets in applications for the automotive market, the results obtained in point positioning remain at the level of several meters. This is mainly due to the previously mentioned design limitations of the smartphones that render phase-based and fixed ambiguity solutions inviable. Without an external antenna, at best ~1m accuracy (95%, GSA, 2019c) (the float solution) can be achieved when using Real Time Kinematic (RTK) corrections (usually linked to extra costs and a need for network coverage). Improvement of the GNSS antenna or using an external one can bring the positioning accuracy with RTK corrections down to sub-decimetre levels (GSA, 2019c).

It is reminded that the position of the device is retrieved from the GNSS antenna and for devices using external antennas, the user guidance should include instructions on the importance of keeping the antenna close to the camera at the photo acquisition time and remaining at the viewpoint some seconds before taking the picture.

4.3 Camera heading accuracy

Similarly to the GNSS chipsets, the choice of the three-axis magnetometers, commonly used in smartphones to determine the orientation relative to the north is a compromise between design constraints, cost of components, power and accuracy (Blum, Greencorn, & Cooperstock, 2013). Compass headings might be influenced by the user's body position, how the phone is held, the environment around the user or even by the device's own magnetic field (e.g. speakers and microphones are made of magnets) that can be also changing with fluctuating processing power. Techniques for filtering the noisy data or for making use of sensor fusion (with a gyroscope or an accelerometer), help to increase the accuracy of camera heading. The estimated average accuracy of the camera heading reported in the literature can be summarised as being close to 10º (Deng, Wang, Hu, & Wu, 2015; Michel, 2017).

In order to maximise the accuracy of the camera heading measurement, the compass/magnetometer of the phone should be kept calibrated. The purpose of calibration is to establish compensation for how the components in the phone (such as screws, speaker magnets, etc.) interfere with the measured magnetic field. The compass calibration is usually performed by repetitive 8-shape movement of the phone, or by rotation around all three axes of the device. During such movements, the magnetometer records changes in the measured field and uses them to calibrate the x, y, and z magnetic field sensors. Requests for periodical sensor calibration should be considered when defining requirements for the app collecting geotagged photos.

4.4 Obtaining the fit-for-purpose picture

It is of high interest to provide sound instructions to the operator/farmer so that administrations will receive and process only useful and fit-for-purpose photos. Such instructions can be provided through dedicated training sessions and/or user guides.

Today capturing geotagged photos is mainly done through the use of mobile applications (App). Support and instructions to capture a specific element can be embedded in dedicated App.

4.4.1 Obtaining the fit-for-purpose picture

Mobile platforms have become an indispensable part of our daily life and routines, but, to date, no standards exist for the design of mobile graphical user interfaces. The principles of desktop software design are not directly applicable, not only due to the limited screen size of mobile devices, but also due to their non-stationary usage, i.e. while walking or driving (Berkman & Hoober, 2011). The industry providing the major operating systems on mobile devices does publish guidelines for how applications should be designed, but the opinions on the best practices differ, both in scientific writing and among industry professionals.

A number of studies have searched for features that make mobile applications usable. According to ISO 9241-11, usability is defined as to which extent a system, product or service can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use (ISO, 2018). Usability is a more precise term than the expression "user-friendly". For users, a usable product should be:

- easy to become familiar with, even during the first contact;
- easy to reach their goal through using it;
- and easy to suppress elements of the user interface for their next usage.

The industry recommendations for developing successful mobile applications (Apple, 2019; Google, 2019) advise keeping consistency across all mobile platforms (uniform elements, spacing and spatial organisation), designing an intuitive user interface and deploying predictable layouts, with legible text regardless of the size of the device screen (Google, 2019). A usable app should behave in a predictable way, i.e. as expected by the users (Apple, 2019). A balance is required between the freedom of user control and the prevention of unwanted outcomes. Ideally, the app should make the users feel they are in control by using familiar and predictable interactive elements, by providing a possibility to confirm destructive actions, or by cancelling operations already in progress (Apple, 2019).

Minimizing actions and screen touches are of major interest for increased usability on mobile platforms. In small handheld devices, without a mouse and a keyboard, it is far more burdensome to select objects and input information than it is on desktop computers. Therefore, scrolling is favoured over clicking and is it advised to optimize a mobile app to keep clicks and field entries to an absolute minimum.

4.4.2 Offline app functionality

Europe has launched a digital programme aiming at, among others, bridging the current digital divide and ensuring everyone has access to networks even in remote areas. For now, still many locations especially in rural areas suffer from unstable or blank Gsm Signal Network. To contrast this issue, app functionalities should be able to work also offline. Thus, supporting data such as the field borders, points from which the photos should be collected, background maps or thematic layers, should be downloaded and stored locally in the mobile device prior to the planned photo capturing. Such data will help guiding the user to the locations where the photos should be taken, display ancillary data (i.e. corresponding LPIS layers, patches of orthoimagery covering the parcels of interest, other thematic layers etc.) and to instruct on how to capture the images.

In cases when orthoimagery is provided as one of the layers displayed in the app, information about the image acquisition date should be provided together with a reminder that the current situation in the field might not be reflected in the image.

4.4.3 Provide user guidance and examples

Due to the multitude of situations where geotagged images are collected (wide variety of landscapes, areas of interest or type of auxiliary data used), it is very difficult to provide universal rules that are applicable in all cases. Nevertheless, it is highly recommended to provide the following minimum set of information to the operator well ahead of the requested photo capturing deadline to minimise his/her efforts and maximise chances to obtain valuable evidence:

- Where the photo has to be captured? Unambiguous indication of the location from where photos should be taken should be provided, e.g. in a form of geographic coordinates of the desired location or a region;
- When should the photos be captured? (E.g. very relevant for automatic crop recognition, when the plant development stage may be linked with the recognition success rate, less relevant for man-made objects that do not evolve dynamically in time);
- What should they depict? (Are multiple objects/subjects of interest allowed in the camera view?);
- How should the frames be taken? (Distance to the object or scale? Camera pointing horizontally? Pointing downwards? From an elevated point? Etc.);
- How many photos should be captured?
- How and when should they be submitted?

Such instructions / guidelines can be conveyed in multitude of ways, e.g. using in-app guidance and tips, tutorial videos, posters, leaflets, written (step-by-step) tutorials etc. It is warmly advised to ensure consultation of end–users to define what would be the most effective methods with the highest acceptance rates in the community of stakeholders.

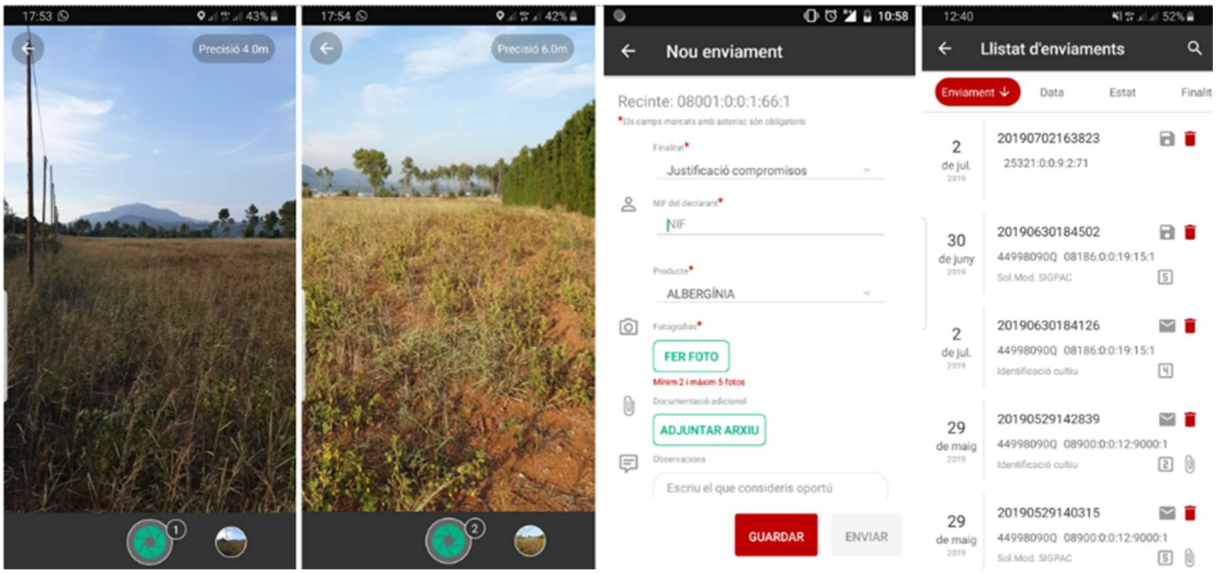
Years around 2018 have seen the onset of smartphone apps development by the Administration of some MSs and by private companies to collect geotagged photos by farmers, farm advisors and/or inspectors. But the huge potential of the geotagged photo collection via a dedicated app has been recognised much earlier and resulted in several very advanced solutions, including the use of augmented reality, being already commercially available and in use in 2018, i.e the ABACO GEOPHOTO, (Abaco, Italy) has been tested and used by the Maltese Paying Agency (ARPA) and the Public Service of Wallonia (SPW) or the e-Geos GeoTag App (Italy) tested by different paying agencies (Austria, Belgium, France, Germany and Spain) and used operatively by the Italian paying agency (AGEA) since 2018. The list of apps availability is currently fast growing (e.g. DIONE and NIVA projects).

Another notable example is the FotoDUN app (DARP, 2019) developed by the Ministry of Agriculture, Livestock, Fisheries and Food (DARP) Government of Catalonia, Spain. This app was made available to farmers to facilitate submitting photos supporting non-conclusive monitoring cases as well as requests for the LPIS update. The development of the app followed a pilot project in which the farmers could upload photos supporting requests for the LPIS data update via an online web form. In both cases, the Web platform and the app, the farmers were provided with a short manual/presentation (digital version, downloadable in a .pdf format) on:

- the link (URL) to the web form or the app installation,
- which credentials to use,
- the data collection protocol,
- the minimum number of images to be taken,
- how to complete the required fields in the photo submission interface,
- examples of good and bad quality photos,
- examples of when the photo evidence should support the claim or support a change of the LPIS.

The app is limited to solely capturing the photos, providing required metadata and submitting them to the Administration server. Prior to going to the field, the background (most recent) orthophoto and the reference data (LPIS and Geo Spatial Application GSA layers) can be downloaded and saved locally to ensure usage even without an internet connection. The user is left free to choose the geographic location and orientation for the photo framing, but recommendations are provided in the user manual. At the time of the photo capture, the user interface indicates the estimated positioning accuracy (photo capture is disabled if the positioning quality is too low) and the number of photos already collected at the location (see **Figure 4**). The fields providing metadata required for photo submission consist of drop-down menus with pre-defined categories, fields to attach photos and other documents and a comment field where free text can be inserted. Photos with the metadata may be sent to the Administration right after capturing or saved in the app for delayed delivery (if required).

Figure 4. Example user interface screens during photo collection with the FotoDUN app developed by the Ministry of Agriculture, Livestock, Fisheries and Food (DARP) Government of Catalonia. A) and B) the user is informed about the number of photos captured at the location. Shooting photos is enabled only if the estimated positioning precision is below a predefined threshold. C) User interface allowing submission of the photos to the System. D) A list of data submitted to the Administration (marked with an envelope icon) or captured and saved for delivery on request (marked with a diskette icon)..



Source: (DARP) Government of Catalonia

Other examples of advanced app functionalities cover, among others:

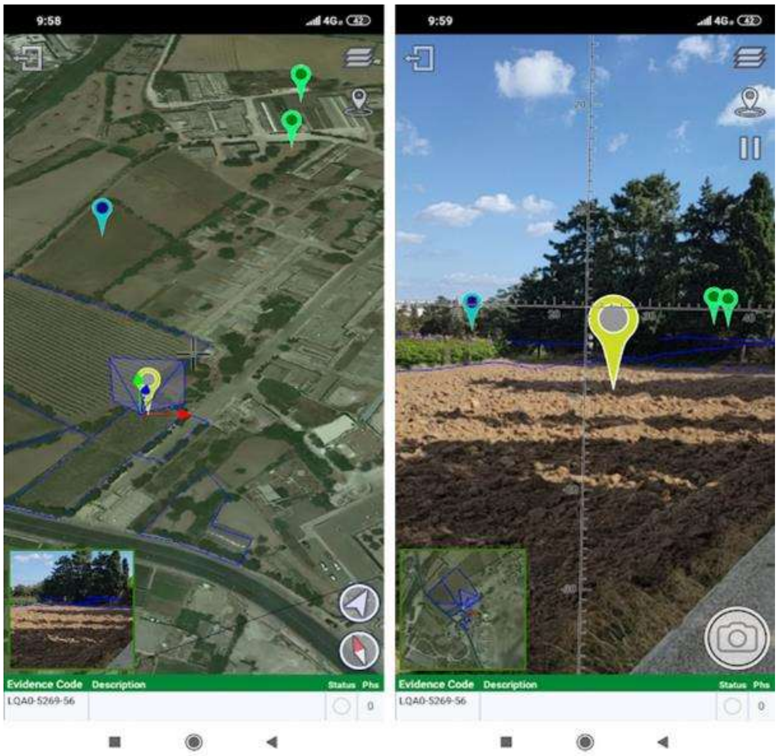
- two-way communication between the Administration and farmers, including the possibility to send a specific task (action request via push notifications) to the farmer or chat with an advisor,
- navigation to the point indicated in the task (action request),
- adapting the information displayed on the screen and the range of allowed actions depending on the proximity to the point (e.g. a photo can only be captured in the vicinity of the destination point, or displaying the "horizon line" when photo capture is allowed) or allowing capture only in a range of vertical and horizontal angles,
- interoperability with many open standards databases, e.g. with existing IACS database to visualise the reference parcels and declared crops (GSAA), ortho-photos, soil samples, other custom indicators or traffic-light provided by the CbM,
- interoperability with Farm Management Information Systems enables access to farm's proprietary data or data collected by in-field sensors, as well as to the logbook of tracked field activities,
- measurements of dimensions and areas both in the geotagged photos and in the 3D view with Augmented Reality,
- visualising IACS information as augmented reality through the mobile camera.

Examples are provided in **Figure 5** to **Figure 7**.

Figure 5. ABACO GEOPHOTO: typical geotagged image acquisition workflow (with augmented reality). After receiving a task, the user is navigated to the point from which the photo should be captured. The current device position and the parcel with the target point are shown on an overview ortho-photo map. The number of photos collected for this task is also made visible. Once the geotagged photo is taken a note can be attached.

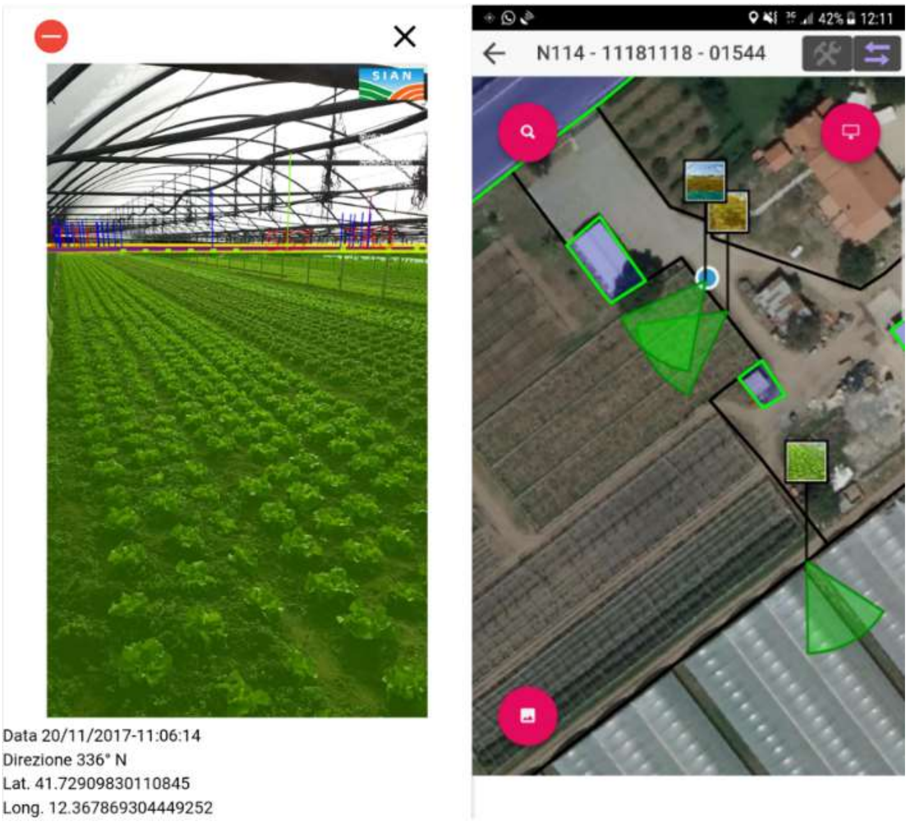
Source: (DARP) ABACO Company, Italy

Figure 6. ABACO GEOPHOTO: augmented reality supporting orientation and data collection. Vertical pin markers indicate parcels and/or features to be captured. The colour of markers indicates the nature/scheme of concern (e.g. yellow: BPS, green: EFA ...). Parcel borders overlaid on the ortho-photo image are displayed in a small window on the screen to ease locating and orientation of the user and photos.



Source: (DARP) ABACO Company, Italy

Figure 7. e-Geos Geotag App: augmented reality supporting orientation and data collection. Vertical blue markers indicate the vertices of the parcels and the nearby parcels. Image location, orientation and thumbnails are shown in the overview.



Source: (DARP) e-Geos Company, Italy

Augmented reality has proven to be a very effective guidance and learning tool in domains like: tourism and navigation, training and education, entertainment and advertisement as well as assembly and maintenance (Chatzopoulos, Bermejo, Huang, & Hui, 2017; Ramos, Trilles, Torres-Sospedra, & Perales, 2018). In mobile augmented reality (often called MAR) wireless communication, location-based computing and services (LBS) and augmented reality are combined to create an integrated interactive environment (Swan, Kuparinen, Rapson, & Sandor, 2017). Although such a combination is still a challenge for user interface designers (e.g. the small display of the mobile devices, potentially low sensor accuracy etc.), MAR is undoubtedly an innovative technology with huge potential in GIS and geospatial tasks (Knowledge@Wharton, 2018; Werner, 2018). Still, it should be used with caution taking into account the hardware limitations of average and lower-end smartphone devices available on the market. Despite the rapid advances, mobile phones' performance as a computing platform for real-time applications is still limited since many technological aspects (e.g. extremely high bandwidth, ultra-low latency and massive connectivity …) remain to be addressed to reach the envisioned future promised Mobile Augmented reality (Siriwardhana *et al.,* 2021).

4.5 Content of user guidance

4.5.1 General considerations

The image capture protocol should be fit for the purpose. The information must be given to the operator such as the trigger for sending the photo (voluntary, compulsory etc.) and the purpose, e.g. confirmation of mowing, confirmation of crop type, request for LPIS update, clarification of irregularities in the crop development (i.e. caused by a flood, drought etc.), data to assess biodiversity or proof of specific plant species etc. Inclusion of this information as photo metadata (or image tags), may prove to be very useful at the stage of image sorting, automated data processing or refining stage of training sets.

Regardless of the intended image content, it is advised to remain at the viewpoint where the photo will be taken for several seconds, and hold the camera motionless before releasing the shutter. This is due to the fact that smartphone providers implement sensor data fusion and smoothing algorithms to determine the geolocation and the compass/camera heading readings. Such algorithms introduce a small delay in the provision of the position/orientation and remaining in the position for a short time before the photo is taken is likely to increase the precision of the geotag photo position and camera heading.

For many situations, it will be advised to capture photos in landscape format (horizontally) and point the camera to depict the element to be evidenced in the image centre. Only relevant objects should be included in the image frame, i.e. fingers, personal belongings, vehicles or other people should not be in the picture. Furthermore, local privacy rules should be respected at all times. Photo anonymization, according to the MS rules and the EU General Data Protection Regulation might be required at some point in the data exchange.

Photos should be captured in good light conditions so that the object is illuminated and clearly visible in the image. Rules for using camera flash should be communicated to the farmer. Photo shooting towards the sun may result in sun glare and cause problems for both automated and visual image analyses and thus should be avoided.

4.5.2 Types of photos

Considering the element to evidence, the operator should be oriented to one or several of three main photo framing categories i.e. “overview”, “close up” and “nadir” photos.

4.5.2.1 Overview photos

An overview photo should depict a large part of the field and include, if possible, landscape elements other than the main object (crop, activity etc.). This type of photo aims at reducing the uncertainty linked with the limited accuracy of the geotagged photo and at providing an overview of the field conditions. The photo should be captured so that a border/corner of a parcel and/or nearby landmarks (trees, ditches etc.) are visible and identifiable in relevant ortho-photo data, thus confirming photo location indicated in the geotag and often for better positioning accuracy. Such photos should be taken with the camera oriented horizontally, with the horizon line falling at approximately 5/6 of the image height to limit the image area depicting the sky (following the image collection protocol used in the LUCAS 2018 survey (E4.LUCAS (ESTAT), 2019)). The object should be centred in the photo frame.

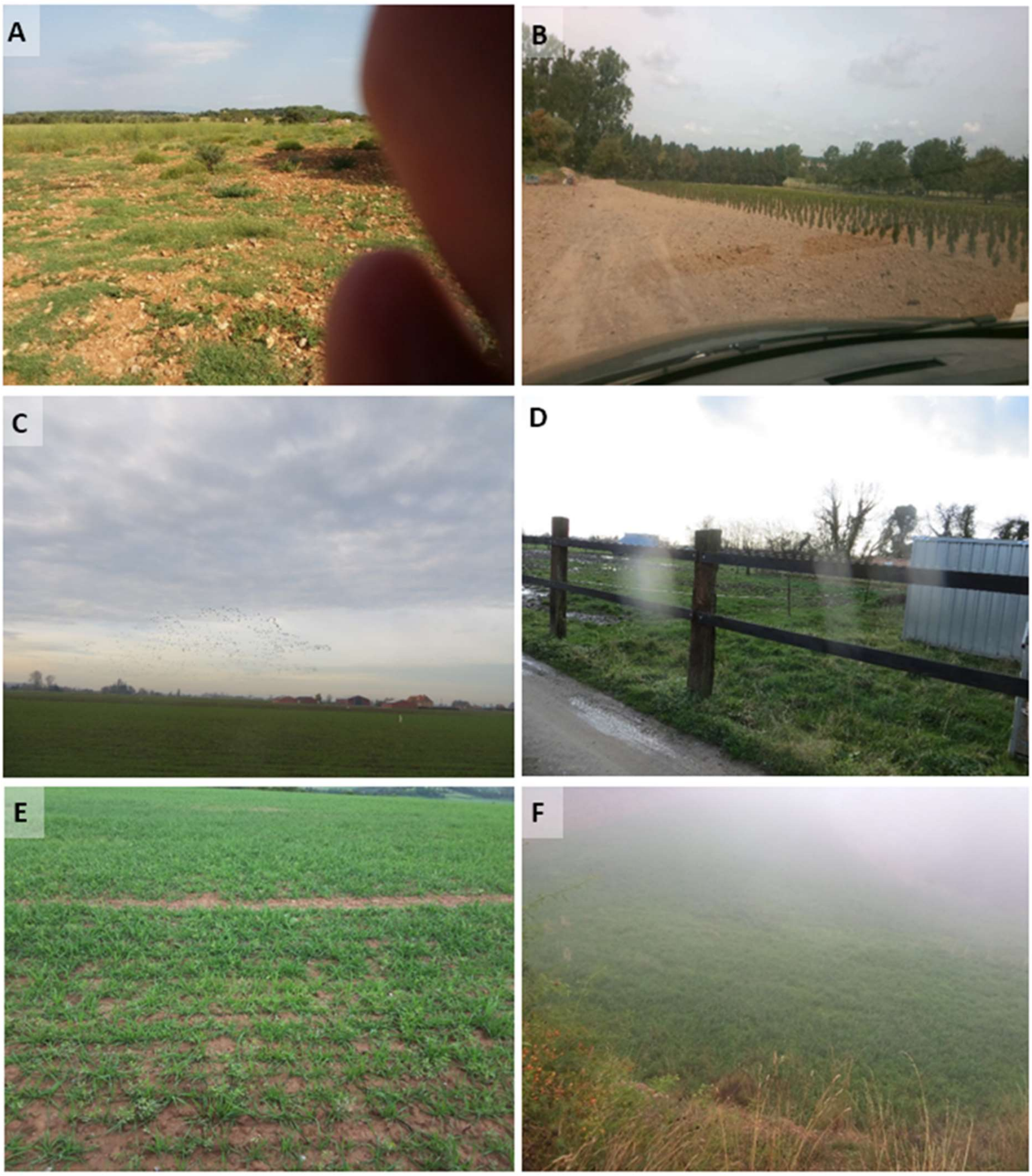
Panoramic photos, being an on-the-fly created mosaic of multiple images, and covering even up to 360 degrees views could be considered as a type of overview photos.

Figure 8 and Figure 9 provide correct and incorrect illustrations of scene “overview” photo capture.

Figure 8. Examples of correctly captured overview photos

Source: Courtesy of the Catalonian Paying Agency, Spain

Figure 9. Examples of overview photos not following the guidelines: A) irrelevant objects in the picture frame, no landscape features visible; B) irrelevant objects (dashboard) in the picture frame, object not centred, camera pointing too high; C) camera pointing too high; D) obstructed view, dirty/wet lens, camera pointing too high; E) camera pointing too low, no landscape features in the photo frame; F) poor visibility, no landscape features in the photo frame.



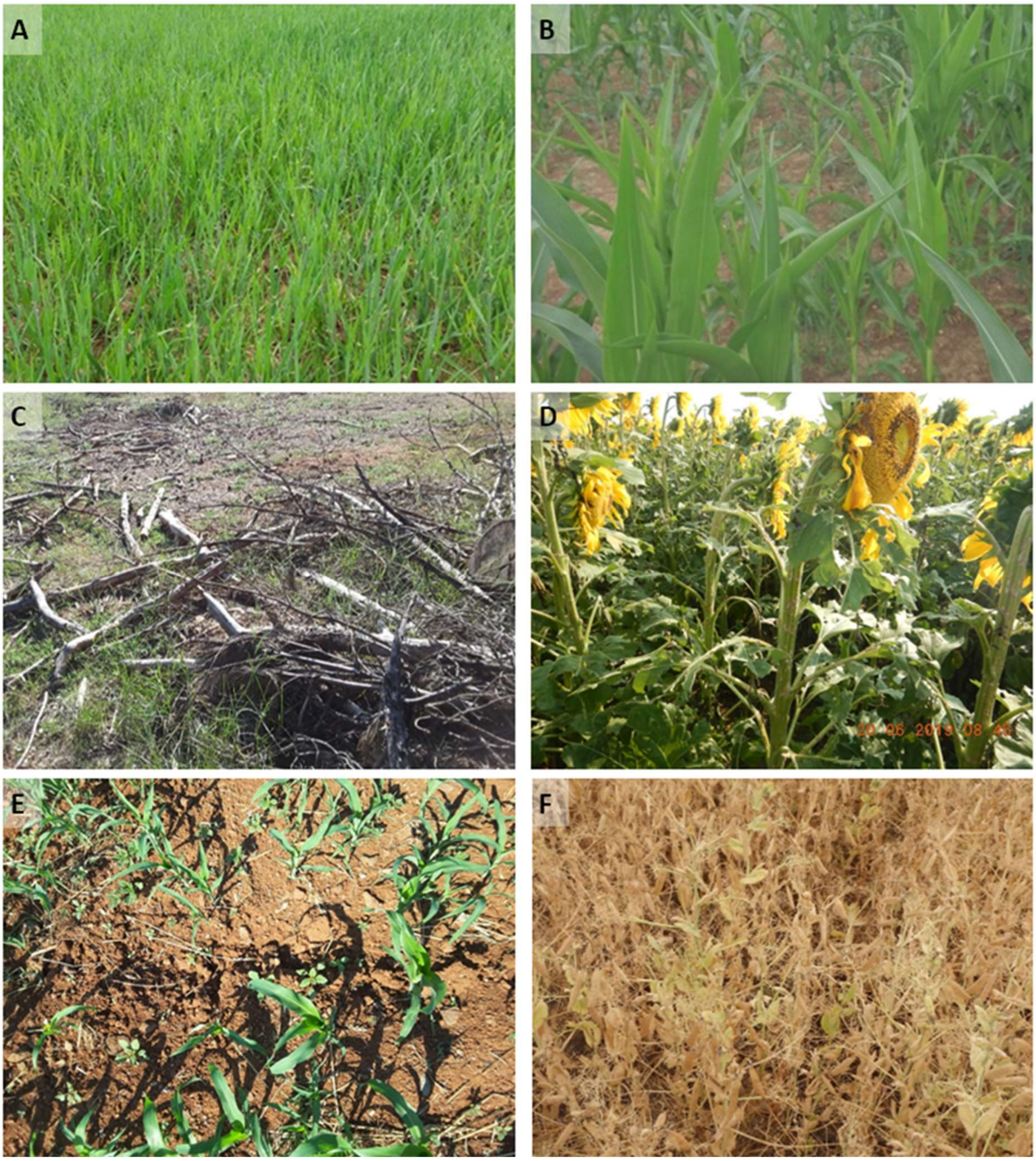
Source: (Photos A, E, F: Courtesy of the Catalonian Paying agency, Spain; Photos B, C, D: Courtesy of the Belgian-Flemish Paying Agency).

4.5.2.2 “Close-up” photos

Due to the fact that overview photos should by nature include a lot of information (mixed landscape), it will be more challenging to handle them (recognise detailed image content) in an automatic way. Therefore capturing a “close-up” photo is advised shortly after (and thus from a nearby position) each overview image, and with an approximately similar camera heading. Such a photo may serve to enable a robust automated identification of the element to evidence. Depending on the characteristics of the plants/objects or/and their development stage it might be advised to take a photo with the camera pointing slightly downwards or horizontally. If a subsequent automatic image analysis step is foreseen, further recommendations could be provided, such as a ban on multiple crops and other objects within the frame or to focus on a specific part of a plant (flower, leaf, fruit …).

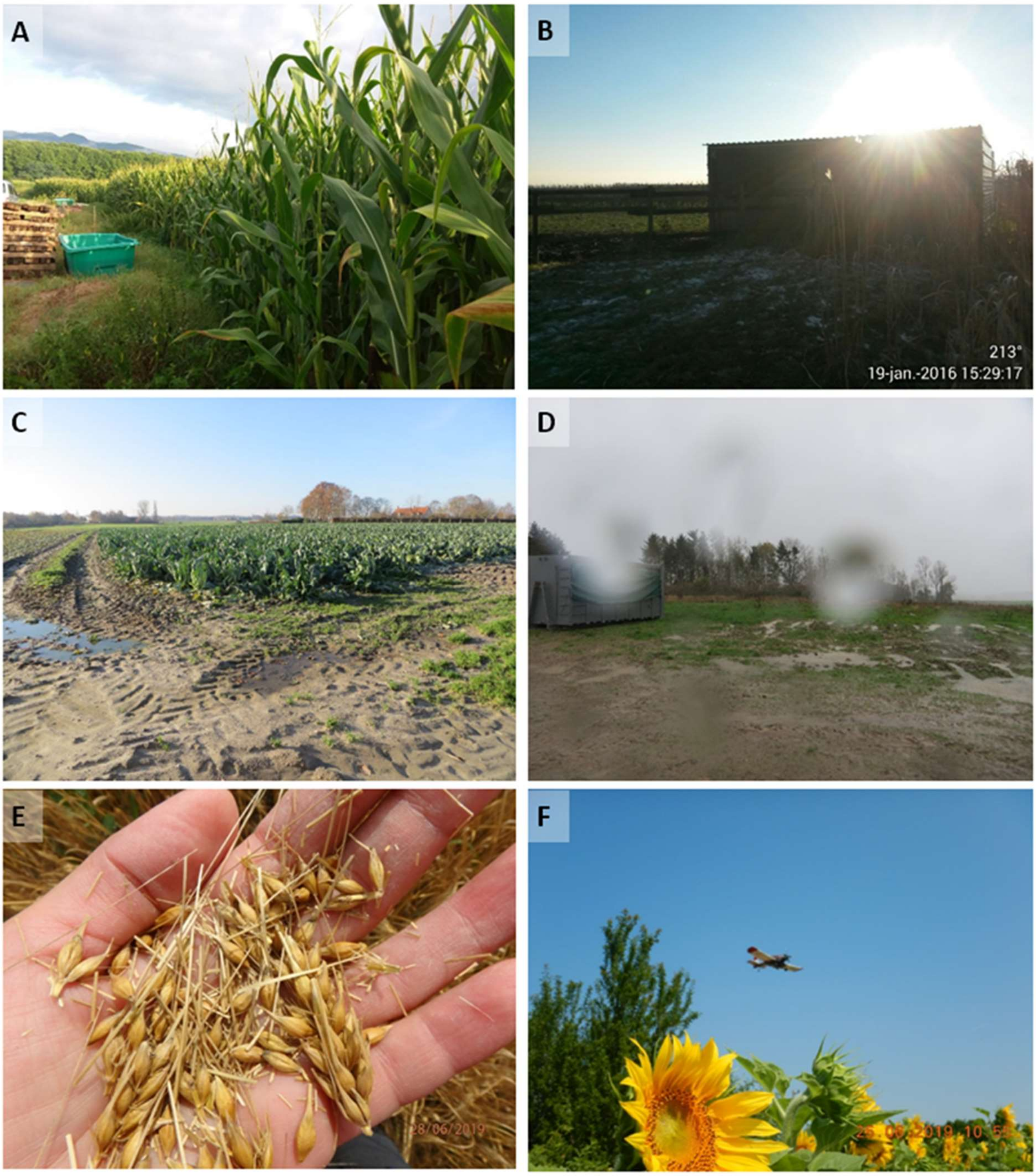
Figure 10 and **Figure 11** provide illustrations of correctly and incorrectly captured "close-up pictures".

Figure 10. Examples of correctly captured close-up photos



Source: (Photos A, C, E: Courtesy of the Catalonian Paying Agency, Spain; Photos: B, D, F: Courtesy of the Bulgarian Paying Agency).

Figure 11. Examples of close-up photos not following the guidelines: A) object not centred; B) photo taken towards the Sun resulting in low contrast; C) Photo taken from a too long distance; D) wet lens decreases the image quality, too long distance to the object; E) irrelevant objects in the image frame (hand); F) irrelevant objects in the image frame, camera pointing too high.

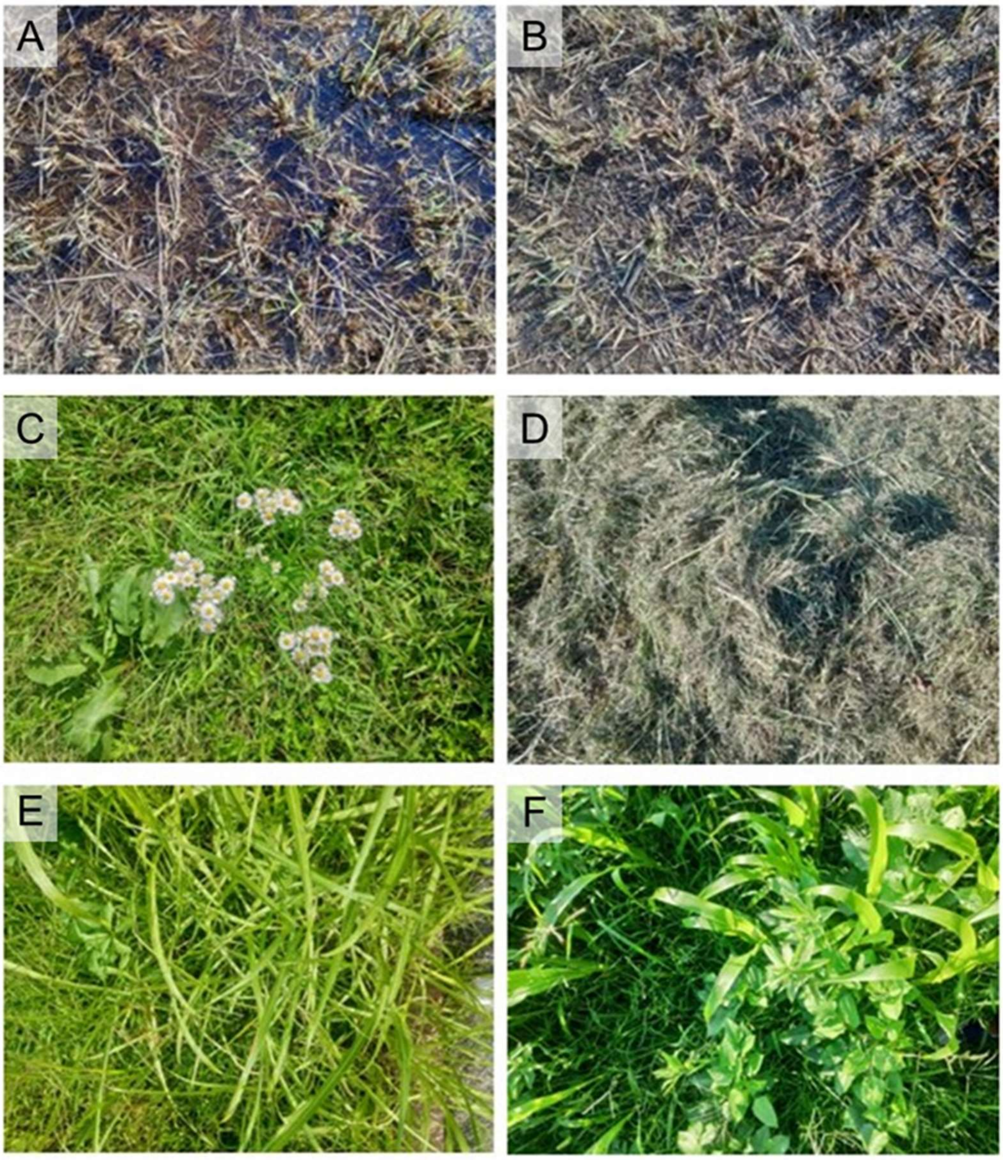


Source: (Photo A: Courtesy of the Catalonian Paying Agency; Photos B, C, D: Courtesy the Belgian-Flemish Paying Agency; Photos E, F: Courtesy of the Bulgarian Paying Agency).

4.5.2.3 Nadir viewing photos

Specific situations may require photos to be taken at a nadir viewing angle (camera pointed vertically downwards). For instance, documenting the crop density, the soil coverage or even the plant richness will be best depicted with nadir photos from an elevated viewpoint. Requesting the photo provider to include the height of the camera in the metadata will allow, when combined with information about the camera matrix and the lens, to derive approximate scale of the objects in the photo.

Figure 12. Examples of correctly captured. A) flooded parcel; B) Parcel A some days after flood; C) Semi-natural pasture; D) Mowed parcel with hay on ground; E) Encroachment; F) Mix cropping.



Source: All photos are courtesy of report's authors

4.5.3 Choice of the appropriate viewpoint and number of images

When requesting geotagged photos from the stakeholder, the intention is to obtain sufficient information to avoid any physical field visit. Therefore, and often, the collected images should provide at least an overview of the parcel, but not necessarily cover its entirety and all details. The total number of photos should depend on the land cover/ activity or object type that needs to be illustrated. Nevertheless, in order to assure a comprehensive view of the element and to limit the possibility of image manipulation, for the overview photos it is recommended to provide at least 2 photos of the element captured from different viewpoints or with different camera headings.

Depending on the purpose of the photo, more or less strict specifications of the photo viewpoint can be provided. In some situations, the operator may be directed to the exact viewpoint and be requested to take a photo with a specific camera heading to ensure that the area of specific interest is depicted in the photos. With the viewpoint and the camera heading pre-set, advice on the optimal time of the day can be provided in order to avoid photos taken towards the sun and thus most probably of insufficient quality.

When leaving the choice of the photo viewpoint to the operator, general instructions should be provided to optimise the result, e.g. to ensure good visibility into the parcel, landscape features on the horizon, adjust the height of the viewpoint to the land cover height etc.

4.5.4 Data integrity and data quality checks

4.5.4.1 Data integrity

In order to ensure the integrity and security of geotagged photos an operating procedure needs to be foreseen for exchange with the administration and safe data storage in the device. Such a procedure should ensure that the information content cannot be falsified at any time, e.g. by geotag manipulation or image content alteration. This task can be addressed in many ways, e.g. by using steganography-based techniques (Mazurczyk & Caviglione, 2015), such as encoding data/codes within EXIF metadata (M. Y. Wu, Hsu, & Lee, 2009) or modification of least significant bits of pixels (Mazurczyk & Caviglione, 2015). In such methods, a code can be encrypted in the photo at the time of image capture and its integrity cross-checked at database entry upon delivery to the Administration. Such measures should also stop attempts to submit images captured in previous years as recently collected data. To detect photo manipulations covering selected image areas algorithms detecting double JPEG compression may be used (e.g. Rahmati et al., 2022).

Another thread to the data integrity is linked to the widespread use of various high-fidelity image displays and printing devices allowing for recapturing high-quality images from screens or printouts. In order to address this problem several recaptured image detection methods have been proposed to date, relying on: complex texture analysis (e.g. Nan and Zhiqin, 2022) or neural networks (e.g. Li et al., 2017). Unfortunately, most of the authors report high algorithm complexity and/or high computational requirements, so their widespread adoption may require more time.

The anti-manipulation measures related to the geotag as such are also available, i.e. checking for installed known applications generating fake GNSS positions, watermarking photos with coded (bar code) location, date and time or cross-checking the GNSS position with the location estimated based on the nearby Wi-Fi and Cell Phone Towers (with compromised accuracy).

4.5.4.2 Image quality

When referring to image quality, the aim is not to obtain aesthetic photos but to get photos interpretable (visually and/or automatically). It means that the photo should not be blurred, too under or overexposed or not containing lens flare.

The human eye has a very high dynamic range (HDR) and also a great ability to focus and adapt to movements, light, depth of field in its field of view. Even best performant cameras are not able to reproduce these human eye capacities. Instructions should be given to ensure a good exposure of photos (e.g. not facing the sun, no dark conditions …) together with some recommendations on the camera setting and use. The 'HDR' setting, available in recent smartphone devices enhances the dark photo areas by adding more details. Since the final photo is a fusion of multiple photos captured with altered camera settings, it requires a still (motion-free) camera position at the shooting moment and will not perform well in very low light conditions.

The object of interest should be in focus in the photo and the user guidance should include a note on that improve the chances of capturing the fit-for-purpose photo (see example **Figure 13**).

Figure 13. Examples of close-up photos illustrating the impact on the choice of the focusing area on the quality of the picture taken.

Source: Courtesy of report's authors

Today the main image file formats are JPEG or JPG files (Joint Photographic Experts Group), TIFF or TIF (Tagged Image File Format) and RAW. JPEG is the far more common file format for photos and often the only format available in digital cameras. However, in JPG format, the image is compressed when recorded. This may induce a loss of image details if the compression level is too high. So, devices should be set to 'Low compression JPG' or 'fine' format to limit distortions (even if the photo size will be bigger). One has to note that most recent cameras allow creating and recording a RAW image and a JPEG image at the same time (RAW+JPG). This setting may be suitable to limit the possibilities of image manipulation. However, since no compression is done on RAW images, their size will often be several times bigger than the JPG equivalent ones. In most areas in Europe, internet and/or GSM bandwidths will be too coarse to allow transferring these RAW format images.

4.6 Image processing

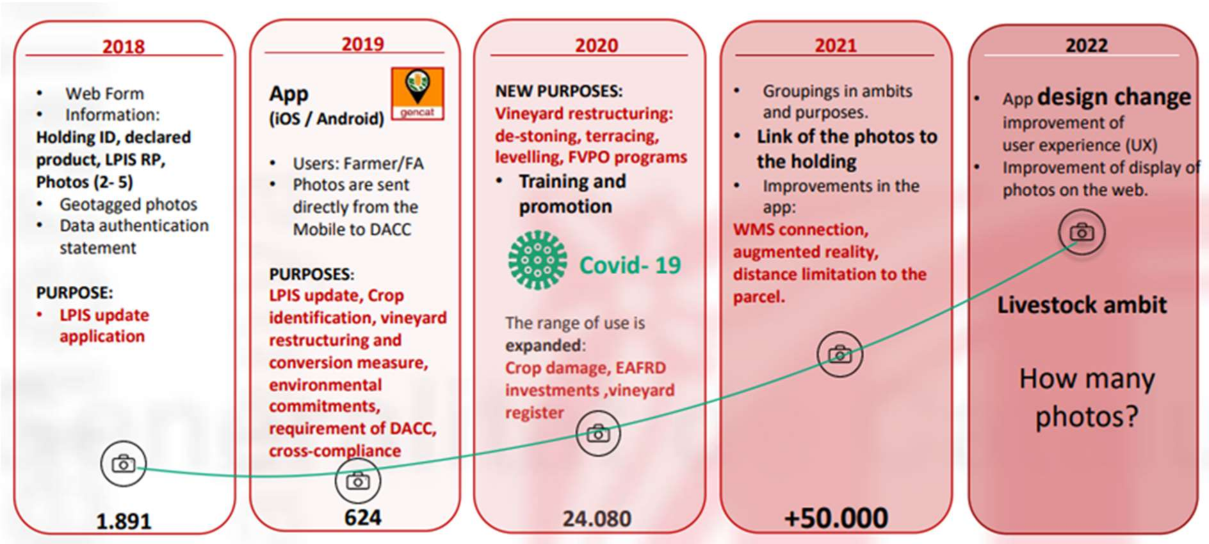
4.6.1 Visual image analyses

Human beings still far outpace computers in their ability to recognise a wide range of photo contents, handling with ease variations in size, colour, orientation, lighting conditions and other factors and this in a fraction of a second.

So for many situations (especially complex patterns) and when the expected number of photos to be received (for a specific element or activity) is low, it will be sufficient and even recommended to check the photos visually.

However, when the use of geotagged imagery starts to be widely deployed in a community of users, the number of data to process can rapidly grow to a state where the visual checking of any single photo will become too time-consuming. For example, the Catalonian Paying Agency started collecting photos from farmers in 2018 (voluntary submission via a web form) and proposed them a geotagged App (PhotoDUN) in 2019. With the successively extending scope of use, in 4 years the number of images collected increased from 1000 to more than 62.000 (in 2022, personal communication).

Figure 14. Evolution of the use of geotagged photos through the PhotoDUN app in Catalonia: from "Experiences of Catalonia on (geotagged) photos provided by farmers: 2018-2021, Paying Agency of Catalonia, ES" presented in GTCAP webinar on geotagged photos and new technologies in CAP, 16 November 2021.



Source: Courtesy of Paying Agency of Catalonia

It is thus recommended that the administration/entity in charge of image processing and management performs relevant cost-efficiency analyses in order to determine the cases and thresholds in numbers for which visual checks become too time consuming and automatic processing should be considered.

4.6.1.1 Define a protocol for visual check and trace activities

Administrations should develop a photo visual check procedure to ensure that all operators will check the same set of predefined elements, trace activities and store results in a standardised manner.

- Data storage

Prior to any data check, the data storage system should be set to avoid modifications or deletions. For instance, geotagged images should be stored as 'read-only'. Also, they should be stored in their original file format. It means that no additional compression process other than the one done at data capture (JEPG) should be performed since critical image information may be lost and artefacts introduced.

The verification procedure may consist of a checklist to verify if a minimum set of information is recorded such as:

- knowing who captured the photo;

The identification of the person who captured the photo should be retrievable. That can be done either through a login-password process and/or through the encoding of the operator identification in the exif file.

- knowing when the image was captured;

The date and time of photo capture should be registered directly through a GNSS antenna (internal, external) of the device. Manual geotagging, by setting date and time directly through the device menu, should be prohibited as it introduces possibilities of error or manipulations. Even if embedded in the metadata file, it is advised to also automatically stamp the date and time on the photo.

Figure 15. Example of stamping of date and time of photo capture directly on the photo.

Source: JRC

- knowing where the image was captured;

It is possible to enter directly the coordinates in the device or by selecting a location from a map using software tools, but this manual geotagging should be avoided.

As for the date and time, the location (longitude, latitude coordinates) of photo capture should be registered directly through a GNSS antenna of the device.

In the case of the overview photos, when they depict a corner of the parcel or an identifiable landmark (e.g. trees, bushes, ponds, fences, characteristic natural landforms etc.) their geolocation can be crosschecked with an up-to-date orthophoto image, if available.

Some other elements can be added to the check list:

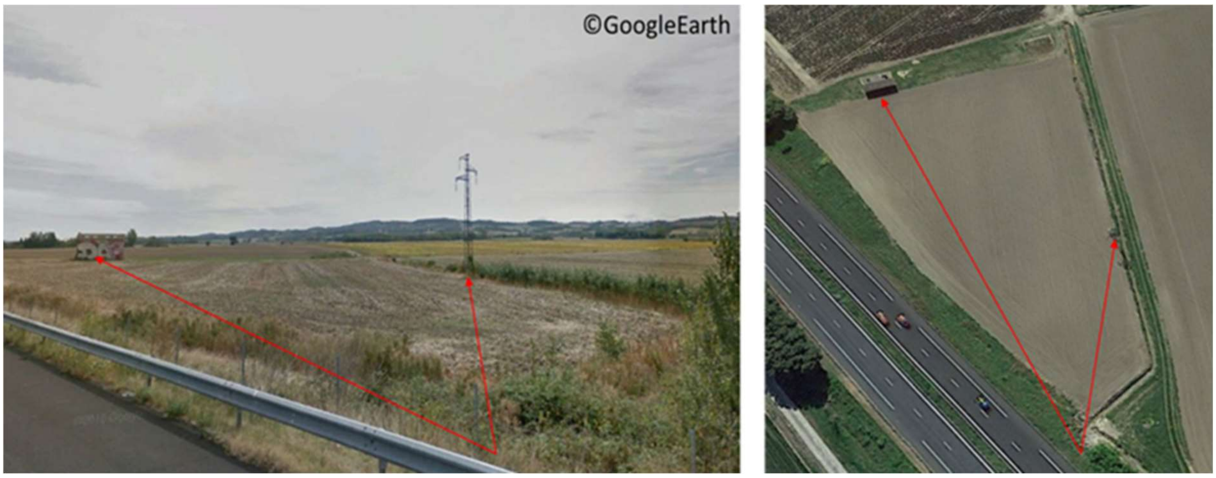
- Checking if looking at an original captured photo (photo integrity);

This step consists in verifying that the whole image information content has, in no manner, been falsified. One common method is to examine the metadata of the image file, and check if information about the image such as the date and time it was taken, the camera model, and other technical details appear to be inconsistent or missing.

- Having information on photo orientation (photo heading);

In addition to the location, it is important to ensure that the camera operator was pointing at the desired object. Crosschecking the approximate camera view indicated by the camera heading encoded in the photo EXIF data with the information in the orthophoto image may provide additional cues on the data consistency. This check is only viable if the smartphone compass has been calibrated prior to image acquisition.

Figure 16. Left: example of a photo containing elements allowing determination of the direction the camera was pointing in when shooting the photo. Right: orthophoto image of the area depicted in the photo. Red arrows point to the same object.



Source: JRC made using Google Earth information as background

- Crosschecking the device technical specification details;

Crosschecking the information linked with the camera model encoded in the EXIF file with the one provided by the manufacturer may provide clues on whether the photo content has not been manipulated.

- Checking the absence of any private information.

It should be ensured that no person or personal information (e.g. car plate number) are recognisable in the photo. In such a situation, faces and areas with private related information should be blurred but, preferably, the photo should be discarded.

Finally, at least two last pieces of information should be recorded.

- The identification of the person who has checked the image together with the date;

- Information on the element to be evidenced and if that evidence has been provided by the photo.

For this last point, it is important to work with a standardised or at least common object terminology/nomenclature. This will ensure better communication between the different stakeholders and a possibility to include the checked image in the training set for automatic photo processing algorithms.

4.6.2 Automated image processing

When the provision of geotagged photos starts to be implemented habitually by beneficiaries, administrations rapidly have to handle and process a very substantial amount of photos, the automatic screening of photos becomes essential.

Achieving diversity in human vision is one of the major challenges for AI research. In the vast majority of cases, we are better than machines at understanding the world around us. But, when trained and targeted to a specific process, the ability of algorithms to detect it may become very replicable, reliable, robust and extremely fast.

Thus, with the recent advances in computer vision and machine learning techniques, image content recognition has become a realistic solution, not only to confirm/contradict alleged crop types/activities depicted in the photo but also to filter out objects that are not of interest for the requested element(s) to evidence. Hereafter, are listed some domains in photo processing where automation can bring substantial benefit.

4.6.2.1 Automatic exposure correction

Exposure is one of the most important factors determining the quality of a photograph. In over-exposed or under-exposed regions, details are lost, colours are washed out, making photo interpretation difficult.

To address this issue, some automatic methods like auto-level stretch and histogram equalization have been proposed. These solutions could be embedded in an App to automatically adjust the photo exposure. In case of necessity, it can be possible to make use of photo editing tools (e.g. Snapseed, Gimp, Photoshop …) to adjust the picture before proceeding to visual interpretation.

It is advised to keep the original image files and perform the enhancement on a copy of the original image.

Figure 17. Automatic photo exposure enhancement using Snapseed app (original photo left, automatically enhanced photo right).

Source: JRC

4.6.2.2 Image integrity

Over the last few years, the development of artificial intelligence technology has allowed proposing solutions to check image integrity much faster than with visual methods.

Examining the image for signs of manipulation or alteration is also called image forensics. Proposed solutions use various elements from camera properties to individual pixels, such as:

- Error Level Analysis: this technique searches for inconsistencies or changes through the detection of compression artefacts in an image;

- Pixel Analysis: by examining the pixel data in an image, algorithms try to detect changes in colour, contrast, or sharpness, which may indicate that the image has been altered;

- Reverse Image Search: this technique involves searching for the image on the internet (could also be in a specific databank) to determine if it has been used or modified in any other context. One can cite the Google Reverse image search, TinEye, Bing image search tools among several others.

The recent introduction of AI systems like DALL·E 2, which can create 'realistic' images on request, may foster the development and improvement of image forensics methods to detect such manipulated pictures.

Other methods are based on digital signature encoding and checks such as:

- Hashing: A unique digital fingerprint is created and added at the image capture. A mathematical algorithm is used to convert the image data into a fixed-length string of characters. This hash value is unique, and any changes made to the image will result in a different hash value. Therefore, by comparing the original hash value with the one of the received image, it is possible to determine if the image has been tampered with.

- Watermarking: Watermarking is a technique of embedding a unique identifier into an image to verify its authenticity. A watermark can be visible or invisible and can include text, logos, or other images.

Watermarking is today commonly used by stock photo agencies to protect their images from unauthorized use or distribution. Watermarking was implemented in the e-Geos Geotag App mentioned in section 4.4.3.

Figure 18. Example of image manipulation detection tested and included in an app developed in the frame of the H2020 DIONE project.

Source: H2020 DIONE project

Figure 19. Example of a manipulation detection test by checking the consistency of focal lens information recorded in the image EXIF file.

Source: Courtesy of Czech Paying Agency - SZIF, CZ

4.6.2.3 Image content

Over the last few years computer vision domain has been dominated by deep learning (Rawat & Wang, 2017) technics which seem to be the most successful methods for image content recognition (Kamilaris & Prenafeta-Boldú, 2018). Image recognition is the process of identifying what an image depicts. That usually is a straightforward task for humans, but for computers that "see" photos as a matrix of numbers, such a task is challenging. Deep learning involves complex data processing and modelling, with the feature extraction performed automatically and represented in a hierarchical way to mimic a human's brain. It processes data hierarchically in separate neuron layers. Deep learning algorithms involve components such as e.g. convolutions, pooling layers, fully connected layers, encoders/decoders or activation functions. Numerous convolutions are performed at selected layers of the network, altering the representations of the learning dataset and thus serving as feature extractors while the pooling layers reduce the data size. They could be seen as filters transforming the input data into another image, highlighting specific image patterns (Kamilaris & Prenafeta-Boldú, 2018), similarly to how the human brain layers perceive visual information, the first convolutional layers extract low-level features, such as edges and blobs, and the later transformations assess the semantic part to the image.

Compared to other classification methods, such an approach to learning features makes the convolutional neural network more robust against changes in illumination, shadows or occluded objects. Because the deep convolutional neural networks find characteristic image features autonomously, less effort is required to process new objects (no need to design feature descriptors). On the other hand, a large amount of training data is required to make the learning process efficient (Ferentinos, 2018).

The learning process for a specific classification task can be facilitated and the required training set size minimized by using transfer learning (Hussain, Bird, & Faria, 2018), a method originating in machine learning, in which a model is trained and developed for one task and is later re-used on a different task. In practice, a pre-trained (trained on a large dataset (millions of images)) model can be used, to solve a problem similar to the one that should be addressed.

Another interesting method for the automated processing of photos is visual object detection which aims to find objects of target classes with precise localization in a given image and assign each object instance a corresponding class label (X. Wu, Sahoo, & Hoi, 2019). Following the remarkable successes of deep learning based image classification, object detection techniques have been advancing very quickly in recent years, but still are more complex and demanding in requirements of the input data and parameters tuning. A comprehensive review of the recent methods can be found in (X. Wu et al., 2019).

Numerous commercial solutions for automated photo classification exist, comprising online platforms for customizable/trainable machine learning image content recognition and cloud storage or APIs offering pre-trained machine learning models that can be re-trained and integrated with the in-house development framework.

As an alternative, various successful and popular pre-trained network architectures exist that could be reused in any in-house developments (Canziani, Paszke, & Culurciello, 2016), e.g. Mobilenet (Howard et al., 2017) AlexNet (Krizhevsky, Sutskever, & Hinton, 2017), VGG (Simonyan & Zisserman, 2015) or Inception-ResNet (Szegedy, Vanhoucke, Ioffe, Shlens, & Wojna, 2016) and other.

Among the most popular frameworks enabling the usage of deep learning are: Keras, PyTorch, TensorFlow, Theano, Caffe, TFLearn, Pylearn2 and the Deep Learning Matlab Toolbox .

Also to date, in the plant domain, plant and plant disease recognition became realistic with a high level of identification accuracy. The number of apps developed is rapidly increasing (e.g. Pl@ntNet, FLORA Incognita, FlowerChecker, PlantSnapp, LeafSnap …). Their strength and performance rely on access to huge image collections mostly fed by crowdsourcing. In the case of more specific applications in the agriculture domain, image collection should be organised to build and train classification algorithms. One can note, for instance, the collaboration to enrich the Pl@ntNet dataset with data and photos coming from the European Union's (EU) Land Use and Coverage Area frame Survey (LUCAS) (Van Der Velde et al., 2023).

In general, the effectiveness and reliability of a trained model is depending on the quality of its training dataset (Boulent, Foucher, Théau, & St-Charles, 2019). To ensure high model performance (and thus generalization) training data of sufficient diversity needs to be provided. The training dataset has to reflect the reality of the operational environment of the tool (Jayme G.A. Barbedo, 2018), e.g. training images for an automatic crop type classification should depict the entire plant (if possible) rather than just a single leaf or fruit, although that could be sufficient for other applications like disease recognition. The difficulty level of distinguishing between different crop types is linked with their phenological development stage. Plants do not only change their leaf shape and proportions but also the entire plant's appearance and colour is changing throughout the season. Therefore two species looking alike at a certain point in time and challenging to distinguish for non-experts, might develop features at a later point in their development cycle that allow for eased separation. This fact should be taken into account not only when planning the collection of photos for the training sample, but also when requesting photos from farmers or providing guidelines on how to collect images.

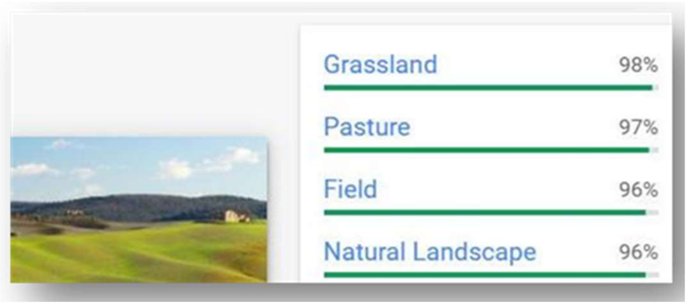
In order to increase the robustness of the classification algorithms, training data collection should be carried out at different times of the season (to cover all development stages of the plant, if necessary), of the day (changing illumination), in varying meteorological conditions and with different cameras (lenses) (Boulent et al., 2019). The variety of maintenance practices should also be covered if these influence the appearance of the plants in images. It is, for instance strongly advised to collect photos of abandoned crops or plants in bad conditions (e.g. after a drought or flooding). As a matter of fact, often geotagged photos will be asked when the automatic processing of satellite data will not be conclusive (the so-called yellow cases). Like in the OTSC method, most of the doubtful parcels in CAPI (Computer Assisted Photo Interpretation) leading to the need for field visits, correspond to 'anomalous' field conditions (due to bad maintenance, the problem of crop development …).

Machine learning methods may be affected by the image "background", especially in complex environments, containing, in addition to the main object/crop, other plants or bare soil (Jayme Garcia Arnal Barbedo, 2016). Nevertheless, the background elements may be linked with the characteristics of certain crops. The models achieve better performance when trained with field images and asked to identify images captured in laboratory conditions (success rates up to almost 68% in Ferentinos, 2018). By contrast, models trained solely with laboratory conditions images performed with a success rate of only 33% in the classification of field images. This proves that also for the machine learning classification of images depicting actual cultivation conditions is a challenging and complex task. It requires a large number of training images to ensure the variety of conditions expected in the images provided by the farmers.

The number of images in all classification classes should be balanced. The number of required images depends on the tools and models used as well as on the classification class complexity and separability. For the purpose of crop/activity recognition, it is advised to use at least several hundred of photos per class. In order to increase the number of training images artificially and to avoid model overfitting, data augmentation techniques can be applied such as histogram equalisation, rotations, image cropping, scaling, mirroring or perspective transformation (Kamilaris & Prenafeta-Boldú, 2018).

The result of the automatic photo classification is usually a list of probabilities of an image belonging to different classes (see example below). The result can be further improved by cross-checking with the metadata provided together with the photo. For example an image submitted as evidence for the cultivation of crop A for which the automatic classification tool assigned class A only as the second highest probability score (so classified as another crop type) could still be accepted as evidence for A if the probability rate of class A for this photo is high enough. Such probability threshold is to be decided through tests performed on the final classification models, trained with the final training sets.

Figure 20. A simulated example of automatic photo classification results. The input image was provided as illustration of a landscape dominated by pasture land. The classification algorithm identified a 'pasture' class in the image with a probability of 97%.



Source: JRC

The performance of classification can be also improved by using specific pre-trained models depending on the type of input photos. For instance, in the Pl@ntNet, when a photo is uploaded, the operator is invited to indicate if the picture corresponds to either a flower, a leaf, a fruit, a bark or the entire plant. Restricting the wide range of possible photo content into more narrow and homogeneous groups allows the usage of algorithms pre-trained on the relevant set of similar photos and results in higher classification accuracy.

Similarly to the quality assurance in the checks by monitoring methodology, the performance of the automated image processing should be quality checked to derive the false positive and false negative classification errors.

Although still technologically demanding or requiring an internet connection, including access to the image classification results in the app for photo collection would bring many benefits. Providing the results of image classification results immediately after the photo capture would reassure the user that the photo captured contains the requested elements and that the photo can be submitted or, on the contrary, it would indicate the need of another photo capture attempt. The submitted photos would be of quality.

4.6.2.4 Image labelling and tagging

The goal of automatic image labelling and tagging is to make it easier for users to search and categorize large collections of images and also to train content-based image retrieval and classification algorithms. Assigning relevant tags/labels by the photo provider or automatically based on the pre-defined purpose of the image acquisition, at the moment of image capture or submission is thus a crucial step.

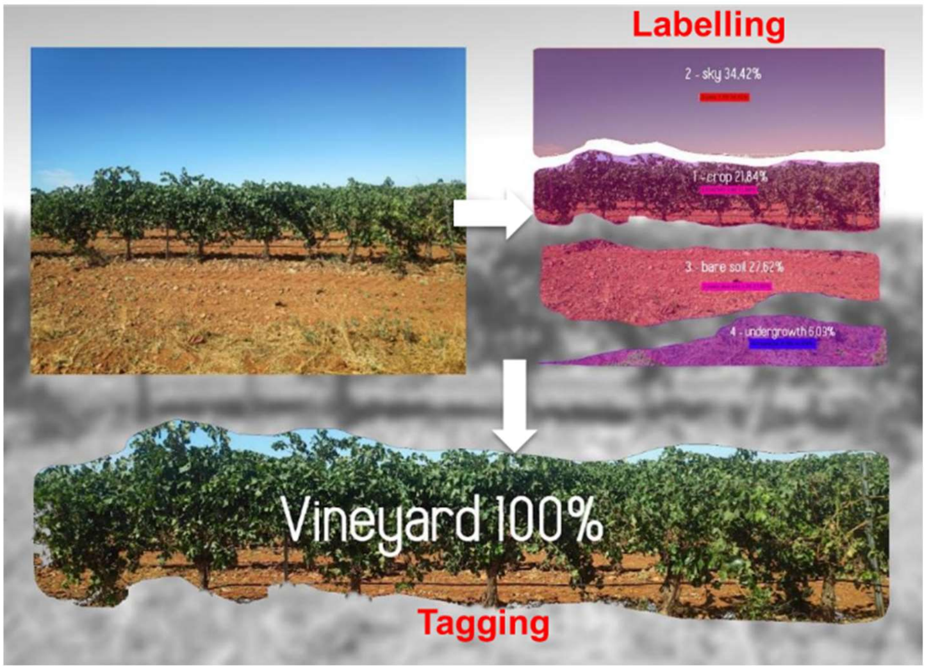
If image labels/tags are unreliable, the training process will not be adequate (Boulent, Foucher, Théau, & St-Charles, 2019) and classification results also unreliable. Thus it is advised to have experts involved at some point of the process to establish a relevant semantic for the catalogue of tags and labels.

The most common approach for automatic tagging/labelling is to use object recognition algorithms to identify objects within an image and then assign relevant tags based on the objects identified with a high level of confidence. (Cf. **Figure 21** and **Figure 22**).

Figure 21. Possible automatic tagging solution. The wild flower (Impatiens genus) is identified on photo with a classification confidence level of 91% (in this case using Pl@ntNet app). 'Impatiens parviflora' and 'Basalm' tags can be automatically assigned to the photo and store in a database. Other possible automatic tags can be attached such as 'invasive species' or 'moist zone'. This will largely ease possible later search, retrieve or further use of the photo.

Source: Pl@ntNet

Figure 22. Automatic segmentation of a photo followed by a set of automatic classifications of each segment content. The tags: 'sky', 'crop', 'bare soil' can be automatically assigned to the picture together with the tag 'vineyard' (SIRIC application, Courtesy of Tragsatec, (ES)).



Source: Tragsatec, ES

4.7 Importance of a synergy between data collection and processing protocols

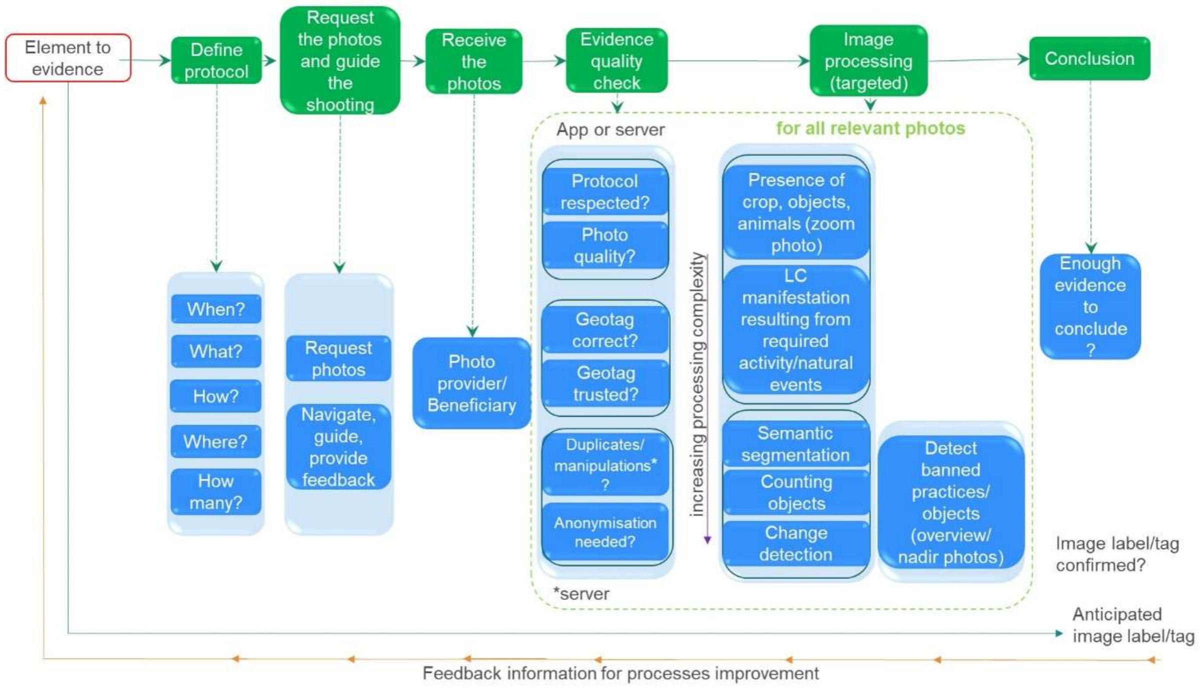
In order to develop efficient and reliable geotagged image capturing solutions, the user information needs and the scenario resulting from the user/farmer declaration should be translated into a set of photo collection and processing protocols.

These protocols should provide clear guidance to the camera operator about when, where and how the photo(s) should be taken so that the depicted image content will be able to serve as an evidence of the required activity and suitable to be processed in an automated manner.

Developing scenarios will consist in defining sequences of stages aiming at optimising what to capture, when to capture it, how to capture it, when to send the request to the person who will take the photo, how to process the photo captured, how to close the scenario. **Figure 23** provides a general workflow of components to consider in the design of the photo collection and processing protocol (based on the more general workflow presented in **Figure 1**).

Figure 23. Proposed general workflow to design scenarios for elements to be evidenced on the ground with geotagged photos.

(Green boxes: workflow steps; Blue boxes: steps components)



Source: JRC

The protocols for collecting evidence, should consider the crop calendar, local practices and other distinctive phenomena (e.g. terrain relief). The data processing protocols will depend of the number of photos to be processed, the advancement of photo processing techniques, as well as the end user's knowledge and experience about local land use and land cover characteristics. The processing solution needs to fit the available budget.

To illustrate these, two examples of data collection protocols and processing workflows are described to evidence a fictitious case of a parcel with high natural value (HNV) grassland.

A 'basic' solution

In the first example (see **Figure 24**), the operators (farmers) population is considered to be highly educated (i.e. very knowledgeable about their farmland and having sound botanical knowledge).

Figure 24. Some components of the fictitious first example of data processing to automatically conclude about presence of the High Natural Value grassland plants on a parcel of interest.

Source: JRC

The scenario will then be that the farmer will be asked to take and provide pictures of specific flowers from his/her field(s).

The associated workflow will then be that:

- At the time of the application for aid, the farmer is reminded that, because he/she declares a HNV grassland parcels, supporting evidence will be needed through the provision of geotagged photos through a dedicated (downloadable) smartphone App.;

- A list of characteristic HNV grassland plants is provided to the farmer (name of plants, photo illustration, main flowering period);.

- The farmer is informed (sms, email, digital dossier …) that he will be instructed that in the flowering period (when flowering plants score the highest plant recognition results in automatic recognition algorithms), he/she will have to take 'macro' or 'nadir' photos (depending on which type of photos were used to train the plant recognition model) on the HNV grassland parcel(s), depicting at least 3 HNV plants from the provided list (each plant in a separate photo);

- During the flowering period, the farmers will go on the considered parcel(s) to take the requested photos with the dedicated App downloaded and installed on the smartphone (and usable through user authentication process);

- When the blooming period starts, a message/sms can be send to the farmer to remind that he/she has to provide photos;

- When the farmer wants to provide the photos. He/she goes to the parcel of interest with a smartphone with the dedicated App installed. (Note that a module can be foreseen to guide the operator to the parcel of interest).

- Once on the parcel the operator is invited to take a close-up photo of one of the flowers of the list provided. When the photo is taken, the App connects automatically to a plant identification app (such as Pl@ntNet or similar). If the plant is identified with a high level of confidence (e.g. >80%) and is part of the HNV plant list, the operator is asked to select the name of flower in the provided list and attached the photo to then send it to the data collecting server (e.g. as part of the IACS). The image is stored and automatically tagged with the Latin plant name and vernacular name.

- Note that a module can be foreseen to authorise the sending of the photo only after check that the geographical position recorded in the exif file is within limits of the parcel of concern.

- A notification in the App informs the user that 2 other different species of flowers still have to be photographed for the considered parcel;

- When 3 photos of different HNV plants (out of official list provided) will be submitted (repeating the previous steps) for a declared HNV parcel, the user will be informed about having completed the task. Completion of such task may last for several days or weeks since the HNV plants present in the parcel may have different flowering periods. Reminders may be set, informing the user/farmer about approaching average beginning and end of the plant flowering periods.

- When 3 photos have been received for a considered parcel, a flag appears automatically in the corresponding farmer's dossier. This information is then used by the administration to further process the dossier.

- In the workflow, since the flowers have been identified automatically with high level of confidence, it can be decided that it is not necessary to visually check all photos received and consider the corresponding parcels as 'compliant' for the HNV request.

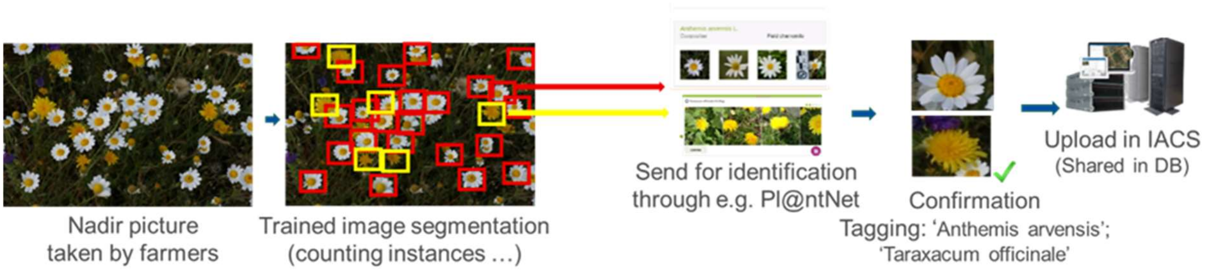
- This should be supported also by a quality check process where a representative number of photos are selected randomly and visually checked to confirm that the image content is compliant with what was automatically identified.

- In case the sample does not pass the QC test, all photos should then be visually verified and action taken to try to improve the plant recognition algorithm and/or adapt the general workflow.

A more advanced solution

In the second example (see **Figure 25**), In this example, thanks to more advanced technological solutions in the photo capturing and processing, the data collection protocol is simpler and no botanical knowledge is expected from the operators (farmers).

Figure 25. Some components of the fictitious second example of data processing to automatically conclude about presence of the High Natural Value grassland plants on a parcel of interest.



Source: JRC

In such case, the associated workflow can vary, compared to the first example, as follow:

- At the time of the application for aid, the farmer is reminded that, since HNV grassland parcels are declared, geotagged photos will be requested as an evidence and should be provided using the dedicated (downloadable) smartphone app.

- At the beginning of the main flowing period (of several HNV plants), a reminder (message/sms) is sent to the farmer about the need to take the photos. The App's advanced user guiding component navigates the operator to the parcel of interest. Once on the parcel the operator is instructed on how to capture an optimal nadir photo depicting several HNV flowering plants;

- Once a photo is taken, the image segmentation is performed and segments are classified to identify individual instances of flowering plants. Each found flower is then sent to a plant identification app (e.g. similar to Pl@ntNet). When at least 3 plants are identified with a high level of confidence (e.g. >80%) and are part of a predefined HNV plant list, the process stops and the photo is authorised to be submitted to the data collecting server. The photo is automatically tagged with the Latin plant names and vernacular name of the plant;

- Eventually the counting instance of each plant can also be registered providing a more detailed indicator of the flower richness of the considered parcel;

- With photos of 3 required HNV plants submitted to the data collecting server, the operator receives a message about a successful completion of the task;

- When photos are uploaded in the database, an operator (from the administration) can see a flag that the task has been concluded positively for the parcel and can further process the dossier.

- Also in that case, since the flowers have been identified automatically with high level of confidence, it can be decided that it is not necessary to visually check all photos received and consider the corresponding parcels as 'compliant' for the HNV request. But an image quality check step should be put in place.

4.8 Quality assurance

As mentioned in the fictitious cases of the previous paragraph, it is highly recommended to include a Quality Assurance step in the general workflow to verify that developed algorithms classify images with a sufficient degree of reliability and consistency.

For this, representative samples of photos that have been automatically accepted as evidence of the presence of a certain land cover or land use should be selected and visually inspected.

The number of samples can be selected based on the number of algorithms developed and used or based on the number of measures/elements to be evidenced.

The visual inspection will consist in verifying that the land cover or land use element for which the photo has been captured is present on the photo. It may also consist in verifying if the labels and tags assigned to the photo (if assigned) are correct.

At the same time, it is recommended to check a minimum set of other elements of the photo and photo processing such as: the consistency of the data recorded in the exif file (date, time, location, camera type …), the identification of the person who captured and transferred the photo.

A comprehensive documentation of the findings should be produced. It should provide at least:

- the numbers of real positive and false positive observed;

- for false positives, the possible reason for the algorithm error like crop confusion, bad image quality (blurred, underexposed …) or bad photo framing (scenery view instead of zoom …).

The results of such analyses may help to improve the algorithm's performance and provide cues for improvements in the data collection protocol or the user in-app guidance.

5 Conclusions

Geotagged imagery offers possibilities to evidence an almost unlimited range of practices, land uses and land covers. Photos can depict a multitude of landscapes, landscape elements or single plants/plant elements.

With the fast pace of technological developments in cameras/smartphones with GNSS antennas, apps dedicated to capturing geotagged photos, plant recognition apps and image content recognition algorithms, handling vast quantities of photos has become realistic.

Organising the sound acquisition and use of geotagged photos holds huge potential for benefit in the frame of check of compliance, knowledge gain, part of the certification process, natural damage evidencing among others and this, for different policy domains (agriculture, environment, climate. …). This is of particular interest for national administrations' to support several of their obligations.

The development of such applications requires a sequence of planned and well-designed investments like: the creation and access to photo collections, the development of ad-hoc image processing and classifiers, the development of apps embedding user-friendly and intuitive solutions to guide operators from data capture up to data delivery to the administration data servers. These steps are not straightforward and can rapidly become highly demanding in terms of time, development and cost.

It is thus of high importance to dedicate time in a preparatory phase to identify the use cases where geotagged photos will offer a clear advantage compared to other solutions (e.g. use of Sentinel temporal series). Then, the element(s) to evidence should be clearly identified and described in the form of scenarios and converted in workflows most adapted to the final users. In case the number of photos to handle becomes significantly high, attention should be paid to the automation of some photo processing phases. Nevertheless, in some situations, the choice of a simple pragmatic solution can be more efficient (in terms of time and cost) than the development of an automated one.

Regardless of the chosen data processing solution, the key to the success is getting the stakeholders/beneficiaries/photo providers actively involved in the process and collaborating. The developed apps and data processing solutions should be endorsed and used by the targeted community. The need for this involvement and providing feedback can be further increased through awareness-raising campaigns (using a variety of communication channels such as emails, newsletters, social media …) and by ensuring training sessions.

Establishment of the best practices or standards for photo tagging and data management in order to maximise the automatic processing efficiency and guarantee high quality results are of high importance as well.

The sharing of experience and good practices is an asset to facilitate all these developments. Such sharing is the main purpose of this report while being aware that many of the underpinned technologies and tools are developing very fast and any guidance will have to be updated accordingly.

References

Apple. (2019). iOS Design Themes. Retrieved July 27, (2018), from https://developer.apple.com/design/human-interface-guidelines/ios/overview/themes/

Ankolekar, A. V., Madappa, R., and Savakis, A., (2022), "Can simpler be better? Review of methods for the detection of GAN-generated imagery", Proc. SPIE 12101, Pattern Recognition and Tracking XXXIII, 121010F; https://doi.org/10.1117/12.2632737

Barbedo, J. G. A. (2016). A review on the main challenges in automatic plant disease identification based on visible range images. *Biosystems Engineering*, *144*, 52–60. https://doi.org/10.1016/j.biosystemseng.2016.01.017

Barbedo, J. G. A. (2018). Factors influencing the use of deep learning for plant disease recognition. *Biosystems Engineering*, *172*, 84–91. https://doi.org/10.1016/j.biosystemseng.2018.05.013

Bekker, A. J., & Goldberger, J. (2016). Training deep neural-networks based on unreliable labels. In *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 20-25 March 2016, Shanghai, China* (pp. 1–10). https://doi.org/10.1109/ICASSP.2016.7472164

Berkman, E., & Hoober, S. (2011). *Designing Mobile Interfaces*. O'Reilly Media.

Blum, J. R., Greencorn, D. G., & Cooperstock, J. R. (2013). Smartphone Sensor Reliability for Augmented Reality Applications. In *International Conference on Mobile and Ubiquitous Systems: Computing, Networking, and Services MobiQuitous 2012: Mobile and Ubiquitous Systems: Computing, Networking, and Services* (pp. 127–138). https://doi.org/10.1007/978-3-642-40238-8_11

Boulent, J., Foucher, S., Théau, J., & St-Charles, P.-L. (2019). Convolutional Neural Networks for the Automatic Identification of Plant Diseases. *Frontiers in Plant Science*, *10*(July). https://doi.org/10.3389/fpls.2019.00941

Canziani, A., Paszke, A., & Culurciello, E. (2016). An Analysis of Deep Neural Network Models for Practical Applications. *Computer Vision and Pattern Recognition*, 1–7. Retrieved from http://arxiv.org/abs/1605.07678

Chatzopoulos, Di., Bermejo, C., Huang, Z., & Hui, P. (2017). Mobile Augmented Reality Survey: From Where We Are to Where We Go. *IEEE Access*. https://doi.org/10.1109/ACCESS.2017.2698164

Dabove, P., Di Pietra, V., and Piras, M., (2020). GNSS Positioning Using Mobile Devices with the Android Operating System. ISPRS International Journal of Geo-Information. 9. 220. 10.3390/ijgi9040220.

DARP. (2019). App FotoDun. Retrieved from http://agricultura.gencat.cat/ca/ambits/desenvolupament-rural/declaracio-unica-agraria/app-fotodun/

Deng, Z. A., Wang, G., Hu, Y., & Wu, D. (2015). Heading estimation for indoor pedestrian navigation using a smartphone in the pocket. *Sensors*, *15*(9), 21518–21536. https://doi.org/10.3390/s150921518

Devos W.,Sima A.and Milenov P., Conceptual basis of checks by monitoring, 2021,JRC127678

Dyrmann, M., Karstoft, H., & Midtiby, H. S. (2016). Plant species classification using deep convolutional neural network. *Biosystems Engineering*, *151*(2005), 72–80. https://doi.org/10.1016/j.biosystemseng.2016.08.024

E4.LUCAS (ESTAT). (2019). *LUCAS2018. Technical reference document C1. Instructions for Surveyors*. Retrieved from https://ec.europa.eu/eurostat/documents/205002/8072634/LUCAS2018-C1-Instructions.pdf

ESTAT. (2019). LUCAS photo viewer. Retrieved from https://ec.europa.eu/eurostat/web/lucas/lucas-photo-viewer

Ferentinos, K. P. (2018). Deep learning models for plant disease detection and diagnosis. *Computers and Electronics in Agriculture*, *145*(September 2017), 311–318. https://doi.org/10.1016/j.compag.2018.01.009

Google. (2019). Material design - Usability. Retrieved July 27, 2018, from https://material.io/design/usability/accessibility.html#

GSA. (2019a). eGNSS4CAP. Retrieved from https://www.egnss4cap.eu/

GSA. (2019b). eGNSS4CAP github repostiroy. Retrieved from https://github.com/SpaceTecPartners/EGNSS4CAP

GSA. (2019c). Third GNSS Raw Measurements Taskforce Workshop, 26th June 2019, Prague. Prague, Czech Republic. Retrieved from https://www.gsa.europa.eu/gnss-raw-measurements-taskforce-workshop

Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., … Adam, H. (2017). MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. *Computer Vision and Pattern Recognition*. Retrieved from http://arxiv.org/abs/1704.04861

Hussain, M., Bird, J. J., & Faria, D. R. (2018). A study on CNN transfer learning for image classification. In *UKCI 2018: 18th Annual UK Workshop on Computational Intelligence. Nottingham, UK* (pp. 191–202). https://doi.org/10.1007/978-3-319-97982-3_16

Iordanov, M., D'Andrimont, R., Lemoine, G., Fasbender, D., Vries, M. de, Voet, P. van der, & Veld, M. van der. (2019). Monitoring Crop Phenology With Street-Level Imagery Using Computer Vision. Retrieved from https://livestream.com/ESA/PhiWeek2019/videos/196125597

ISO. (2018). ISO 9241-11:2018 Ergonomics of human-system interaction — Part 11: Usability: Definitions and concepts quality models.

Kamilaris, A., & Prenafeta-Boldú, F. X. (2018). Deep learning in agriculture: A survey. *Computers and Electronics in Agriculture*, *147*(February), 70–90. https://doi.org/10.1016/j.compag.2018.02.016

Knowledge@Wharton. (2018). Why Augmented Reality Is the Next Big Innovation in Mapping. Retrieved November 7, 2019, from https://knowledge.wharton.upenn.edu/article/whats-next-for-the-innovation-behind-google-maps/

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). ImageNet Classification with Deep Convolutional Neural Networks. *Communications of the ACM*, *60*(6), 84–90. https://doi.org/10.1201/9781420010749

Kučinskas, A. (2021). „NMA AGRO" from geotagg photos to open data. Presentation in GTCAP webinar on geotagged photos and new technologies in CAP, video meeting, 16 November 2021. https://marswiki.jrc.ec.europa.eu/wikicap/images/1/10/06_NMA_AGROdevelopment.pdf

Li, H., Wang, S., Kot, A.C. Image recapture detection with convolutional and recurrent neural networks, 2017. Electronic Imaging, Media Watermarking, Security, and Forensics 2017, pp. 87–91. DOI : 10.2352/ISSN.2470-1173.2017.7.MWSF-329

Nan, Z., Zhigin, L., Recaptured image forensics based on local ternary count of high order prediction error, 2022. Signal Processing: Image Communication, Vol. 104, 2022, pp. 1-10. https://doi.org/10.1016/j.image.2022.116662

Mazurczyk, W., & Caviglione, L. (2015). Steganography in Modern Smartphones and Mitigation Techniques. *IEEE Communications Surveys and Tutorials*, *17*(1), 334–357. https://doi.org/10.1109/COMST.2014.2350994

Michel, T. (2017). *On Mobile Augmented Reality Applications based on Geolocation*. Université Grenoble Alpes. Retrieved from https://hal.inria.fr/tel-01651589/document

Musulin, I., Kos, S., & Brčić, D. (2014). A study of smartphone satellite positioning performance at sea using GPS and GLONASS systems. In *Proceedings of the 22nd International Symposium on Electronics in Transport* . (pp. 1–7). Electrotechnical Association of Slovenia. https://doi.org/10.13140/2.1.3593.7929

Rahmati, M., Razzazi, F., Behrad, A. Double JPEG compression detection and localization based on convolutional auto-encoder for image content removal, 2022. Digital Signal Processing, Vol. 123, pp. 1-14. https://doi.org/10.1016/j.dsp.2022.103429

Ramos, F., Trilles, S., Torres-Sospedra, J., & Perales, F. J. (2018). New trends in using augmented reality apps for smart city contexts. *ISPRS International Journal of Geo-Information*, *7*(12). https://doi.org/10.3390/ijgi7120478

Rawat, W., & Wang, Z. (2017). Deep convolutional neural networks for image classification: a comprehensive review. *Neural Computation*, *29*, 2352–2449. https://doi.org/10.1162/NECO_a_00990

Roberts, W., Critchley-Marrows, J., Fortunato, M., Ivanovici, M., Callewaert, K., Tavares, T., … Pomies, A. (2018). FLAMINGO – Fulfilling enhanced location accuracy in the mass-market through initial GalileO services. *Proceedings of the 31st International Technical Meeting of the Satellite Division of the Institute of Navigation, ION GNSS+ 2018*, 489–502. https://doi.org/10.33012/2018.15931

Robustelli, U., Baiocchi, V., & Pugliano, G. (2019). Assessment of Dual Frequency GNSS Observations from a Xiaomi Mi 8 Android Smartphone and Positioning Performance Analysis. *Electronics*, *8*(9), 1–16. https://doi.org/10.3390/electronics8010091

Seyyedhasani, H., Dvorak, J. S., Sama, M. P., & Stombaugh, T. S. (2016). Mobile Device-Based Location Services Accuracy. *Applied Engineering in Agriculture*, *32*(5), 539–547. Retrieved from https://uknowledge.uky.edu/bae_facpub/16

Simonyan, K., & Zisserman, A. (2015). Very deep convolutional networks for large-scale image recognition. *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, 1–14.

Siriwardhana, Y., Porambage, P., Liyanage, M. and Ylianttila, M. (2021) "A Survey on Mobile Augmented Reality With 5G Mobile Edge Computing: Architectures, Applications, and Technical Aspects," in IEEE Communications Surveys & Tutorials, vol. 23, no. 2, pp. 1160-1192, Secondquarter 2021, doi: 10.1109/COMST.2021.3061981.

Specht, C., Dabrowski, P. S., Pawelski, J., Specht, M., & Szot, T. (2019). Comparative analysis of positioning accuracy of GNSS receivers of Samsung Galaxy smartphones in marine dynamic measurements. *Advances in Space Research*, *63*(9), 3018–3028. https://doi.org/10.1016/j.asr.2018.05.019

Swan, J. E., Kuparinen, L., Rapson, S., & Sandor, C. (2017). Visually Perceived Distance Judgments: Tablet-Based Augmented Reality Versus the Real World. *International Journal of Human-Computer Interaction*, *33*(7), 576–591. https://doi.org/10.1080/10447318.2016.1265783

Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2016). Rethinking the Inception Architecture for Computer Vision. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, *2016–Decem*, 2818–2826. https://doi.org/10.1109/CVPR.2016.308

Tomaštík, J., Tomaštík, J. J., Saloň, Š., & Piroh, R. (2016). Horizontal accuracy and applicability of smartphone GNSS positioning in forests. *Forestry: An International Journal of Forest Research*, *90*(2), 187–198. https://doi.org/10.1093/forestry/cpw031

Tanaka, M.; Shiota, S.; Kiya, H. A Detection Method of Operated Fake-Images Using Robust Hashing. J. (2021) Imaging, 7, 134. https://doi.org/10.3390/jimaging7080134

Van Der Velde, M., Goeau, H., Bonnet, P., D`andrimont, R., Yordanov, M., Affouard, A., Claverie, M., Czucz, B., Elvekjaer, N., Martinez Sanchez, L., Rotllan Puig, X., Sima, A., Verhegghen, A. and Joly, A., (2023) Pl@ntNet Crops: merging citizen science observations and structured survey data to improve crop recognition for agri-food-environment applications, ENVIRONMENTAL RESEARCH LETTERS, ISSN 1748-9326, 18 (2), p. 025005, JRC132465.

Werner, P. A. (2018). Review of implementation of augmented reality into the georeferenced analogue and digital maps and images. *Information*, *10*(1). https://doi.org/10.3390/info10010012

Wu, M. Y., Hsu, C. C., & Lee, J. H. (2009). A GeoTagging scheme using image steganography and GPS information authentication. In *IIH-MSP 2009 - 2009 5th International Conference on Intelligent Information Hiding and Multimedia Signal Processing* (pp. 1245–1248). IEEE. https://doi.org/10.1109/IIH-MSP.2009.301

Wu, X., Sahoo, D., & Hoi, S. C. H. (2019). Recent Advances in Deep Learning for Object Detection. *Computer Vision and Pattern Recognition*, 1–40. Retrieved from http://arxiv.org/abs/1908.03673

Yalcin, H., & Razavi, S. (2016). Plant classification using convolutional neural networks. In *2016 5th International Conference on Agro-Geoinformatics, Agro-Geoinformatics 2016* (pp. 1–5). IEEE. https://doi.org/10.1109/Agro-Geoinformatics.2016.7577698

Zhang, X., Tao, X., Zhu, F., Shi, X., & Wang, F. (2018). Quality assessment of GNSS observations from an Android N smartphone and positioning performance analysis using time- differenced filtering approach. *GPS Solutions*, *22*(3), 1–11. https://doi.org/10.1007/s10291-018-0736-8

Plant recognition App cited:

FlowerChecker App:	https://www.flowerchecker.com/
Leafsnap App:	http://leafsnap.com/
PlantNet App:	https://plantnet.org/
PlantSnap App:	https://www.plantsnap.com/

Deep learning tools cited:

Caffe:	https://caffe.berkeleyvision.org/
Hugging face:	https://huggingface.co/
Keras:	https://www.tensorflow.org/guide/keras
Pylearn2:	http://www.deeplearning.net/software/pylearn2/
Pytorch:	https://pytorch.org/
Tensorflow	https://www.tensorflow.org/
Tflearn:	http://tflearn.org/
Theano:	https://en.wikipedia.org/wiki/Theano_(software)

List of abbreviations and definitions

AMS Area Monitoring System

CAP: Common Agriculture Policy

CAPI: Computer Assisted Photo Interpretation

CbM: Checks by Monitoring

CTS: Common Technical Specifications

EFA: Ecological Focus Area

GSA: European GNSS Agency

GSAA: Geo Spatial Aid Application

GNSS: Global Navigation Satellite System (This term includes the GPS, GLONASS, Galileo, Beidou and other regional systems)

HAS: High Accuracy Service

IACS: Integrated Administration and Control System

LBS: Location-Based computing and Services

LPIS: Land Parcel Identification System

MAR: Mobile Augmented Reality

MS: Member State

OS-NMA: Open Service Navigation Message Authentication

OTSC: On-The-Spot Checks

RFV: Rapid Field Visit

RTK: Real Time Kinematic

2DRMSE 2-Dimensional Root Mean Square Error

List of figures

List of tables

Science for policy

The Joint Research Centre (JRC) provides independent, evidence-based knowledge and science, supporting EU policies to positively impact society

EU Science Hub
joint-research-centre.ec.europa.eu

@EU_ScienceHub

EU Science Hub - Joint Research Centre

EU Science, Research and Innovation

EU Science Hub

@eu_science