CURRENT CHALLENGES IN OPERATIONAL VERY HIGH RESOLUTION LAND-COVER MAPPING

Clément Mallet,* Arnaud Le Bris

Univ. Gustave Eiffel, IGN-ENSG, LaSTIG, Saint-Mande, France
firstname.lastname@ign.fr

KEY WORDS: review, land-cover, geospatial images, mapping, operational, very high spatial resolution, remote sensing.

ABSTRACT:

Many land-cover products have been made available for a large range of end-users over the last ten years, even at global scales. In particular, remote sensing data analysis has proved to be the most feasible solution for automation purposes, at multiple spatial scales. However, current solutions are not sufficient for designing better products, adapted to real-case applications, operational constraints, and the generation of services, built upon these core layers. In this paper, we review the main requirements and the recent changes in remote sensing for the specific case of very high resolution land-cover mapping. We also comment current and evaluate challenges for the optimal exploitation of Earth Observation images with the aim of automatically generating maps tailored to specific end-users' needs. We advocate for more challenging large-scale benchmarks and for human-in-the-loop solutions.

1. INTRODUCTION

Land-cover (LC) can be defined as the observed physical cover on the Earth surface (Giri, 2012). Objects and surfaces are segmented into classes. Their number, type, and definition (integrated into the concept of *nomenclature*) vary with the application and the geographical scale. At global levels, general classes with a coarse spatial division are available (e.g., urban areas, forests, water, crops). At local scales, fine-grained classes become available (building types, roads, tree species). LC description is the core information layer for a large variety of interdisciplinary scientific studies (Pereira et al., 2013). Accurate up-to-date maps over large areas are mandatory baselines. A large number of public policies are driven by such knowledge: climate change mitigation, reduction of risks and threats, global sustainable development (Feddema et al., 2005).

Remote sensing (RS) through automatic image analysis of Earth-Observation (EO) has been widely recognized as the most feasible approach to derive LC information over large areas (Manakos, Braun, 2014). It is now accepted that manual or semi-automatic generation of LC geodatabases through visual inspection of EO images is not a sustainable answer for current needs (Grekousis et al., 2015). Such issue is exacerbated with the increasing demand for semantic and geometric accuracy, and up-to-date knowledge.

However, currently, none of the existing products meet all the required needs. Global databases (E.U., USA, China) exhibit limited semantic accuracy and heterogeneous quality. They are updated at best every 5 years (Yu et al., 2014, Ban et al., 2015, Skidmore et al., 2015). Beyond large-scale monitoring purposes, they are of low relevance. Without flexibility and customization capacities, their exploitation potential is significantly narrowed down. Conversely, local geodatabases can be semantically very detailed but will cover a limited area. They correspond to very peculiar needs. The human effort spent for such a generation can hardly be repeated without significant financial inputs and maps are often obsolete when the collection process ends. Existing solutions have been recently improved with the unprecedented amount of remote sensing images of heterogeneous physical nature resulting from new EO satellites

with short revisit time (Sentinel programme, 1 image/5 days) and increased spatial resolutions. Such data abundance indeed brings new opportunities for the long-term and yearly description of Earth surface, up to the scale of a continent. Coupled with existing Very High spatial Resolution (VHR) data, almost all needs for Earth surface description can be fulfilled provided they can be optimally processed by appropriate algorithms.

In this paper, we briefly review and discuss the status, requirements and the main challenges in generating relevant land-cover maps on a timely manner, applied to operational constraints, for very high spatial resolution purposes (<10 m). Operational land-cover mapping is first defined (Section 2). Then, the paradigm change in remote sensing is presented (Section 3), before discussing current challenges (Section 4).

2. OPERATIONAL LAND-COVER MAPPING

2.1 Current products

Initiatives are increasingly being taken for both land-cover status and change monitoring. Geodatabases and products now exist at various scales, corresponding to specific policies' needs. In general, the finer the spatial resolution, the more targeted the nomenclature. The existing LC databases can be coarsely categorized into four main groups, as detailed below. They can be evaluated through four main features (Figure 1): spatial coverage, spatial and semantic accuracies (linked to the temporal, spatial and spectral resolutions of the sensor of interest), and updateness. They are further described in Section 2.2.

Global maps. They aim to fully described the Earth's surface, such as Global Land Cover Characterization (USGS), Glob-Cover (ESA), or GlobLand30 (China). From >1 km products, we recently moved to 30 m resolution. First products were derived from time series optical satellite data at coarse spatial resolution (300 m-1 km). They did not provide sufficient thematic detail or change information for global change studies and resource management (Bontemps et al., 2012, Giri et al., 2013). Higher spatial resolution and more frequent data products were then developed (30 m products, the spatial resolution of Landsat 7-8 images). With superior details and finer-grained cat-
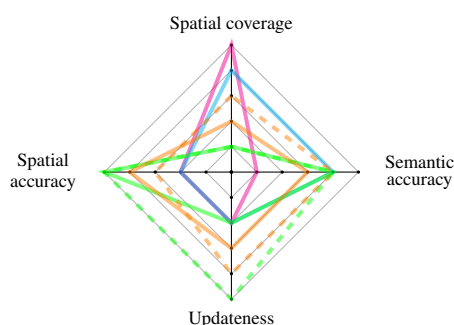
---

*Corresponding author

Figure 1. Categorization and evaluation of the four main groups of land-cover products, with respect to four main features: global, continental, national, local.

egorization, this opens the field for change detection of most human activity and offers the increased flexibility of environmental models needed for global change studies (Buchanan et al., 2009, Pereira et al., 2013, Skidmore et al., 2015).

Continental-scale products. Maps such as the U.S. National Land Cover Database (Homer et al., 2015, Yang et al., 2018), or Copernicus Corine Land Cover (Aschbacher, Milagro-Pérez, 2012) are available at ∼ 25 m resolution (several epochs), and accompanied with change maps (Figure 1). Mostly produced with automatic classification tools, they are complemented with human intervention (visual inspection of RS images). Although global maps are rather general, targeted (class-specific) by-products also exist at such scales: e.g., the tree canopy and imperviousness layers (NLCD), High-Resolution urban layers (Urban Atlas) or the delineation of the riparian areas in the European Environment Agency Copernicus catalogue.

National datasets. They frequently either correspond to standard land-cover maps (full partition on the space) or two-dimensional topographic databases that exist in many developed countries. They have often been produced and updated by national mapping and cadastral agencies and focus on specific objects (buildings, roads, crops). LC maps are often semantically less accurate than topographic databases. The nomenclature often gathers land-cover and land-use classes or proposes only coarse classes (e.g., urban areas). This is sufficient for statistical but not for mapping purposes. Newly (semi) automatic solutions have been recently proposed (Inglada et al., 2017, Jalal et al., 2019) and pave the way for products with superior spatial coverage and updateness (dashed line, Figure 1).

Local initiatives. Two possibilities co-exist. On the one hand, such DBs can be a subset of national products with a higher spatial accuracy and often a higher completeness in terms of objects. This is due to the fact that more frequent updating is possible at such small scales (dashed line Figure 1). On the other hand, they may focus on classes that are not present in topographic databases at coarser scales. They are often generated through initiatives of local authorities.

Automation was not considered here. Fully automatic solutions based on EO image analysis exist, as well as those integrating human inputs or existing land-cover databases. While automation is often inversely correlated to the spatial coverage of the LC DB, the different solutions and products may be hardly comparable: these pipelines are rarely fully described.

In-depth analysis of global land-cover products is available in the literature (Yu et al., 2014), featuring more comparison criteria (sensor, accuracy, classification method, ability to detect

changes). Such papers focus on global scales (Tchuenté et al., 2011, Congalton et al., 2014, Tsendbazar et al., 2016).

2.2 Current requirements

First initiatives in map generation from remote sensing data have highly stimulated various stakeholders and end-users, in recent years. This allows today a sharper definition of the current expectations of the communities in terms of land-cover products and topographic databases.

Frequent updating, related to the capacity in detecting changes, requires techniques to spot specific patterns or update LC DBs established with a significant contribution of manual operators (over several decades). Volunteered Geographic Information (VGI) solutions have recently gained significant interest so as to reduce the manual workload and upscale the updating processes (See et al., 2015, Brovelli et al., 2018). Secondly, in case of missing reference data, post-classification strategies prevail (Tewkesbury et al., 2015). Whatever the strategy, LC should be provided at a regular basis (3→10 years), if possible with the trajectory of changes. However, the frequency may no longer be sufficient to meet monitoring requirements. 2-5 year is now a time lapse explicited by many end-users (it can be lower, e.g., 1 year, for very specific phenomena such such urban changes). The cycle length is often explained by the tedious and time-consuming validation process.

Spatial improvement. It is now possible to sharply delineate objects that were only coarsely detected with first satellites an to improve the planimetric accuracy of topographic objects and LC classes. With steadily increasing spatial resolutions, the field towards VHR mapping is open. The underlying issue is the ability of new RS processes to take benefit from the richness of higher spatial resolutions while keeping the advantages of former sources (temporal and spectral resolutions).

Semantic improvement. Similarly, with increasing spatial, temporal and spectral capabilities of multi- and hyper-spectral RS sensors, new classes can be discriminated and detected. It results in richer nomenclatures (hedges/bushes, roads↔buildings, crop cultures, tree species, etc.), and opens the field to new thematic applications.

Scalability and completion. It is often required to increase the spatial extent of LC maps (=*upscaling*) for two main reasons. First, they may have been created in a limited portion of a territory for some reason (political, financial, non availability of RS source to support their generation at this moment, on-going generation process that needs to be fastened, etc.). Secondly, the spatial coverage may be satisfactory, but completeness issues may exist: (i) some elements of a class may be missing, resulting in an heterogeneous DB quality or (ii) some classes may be (highly) under-represented if the LC DBs were mainly generated from one or several target-specific classification methods. In both cases, all the classes of the nomenclature are known but with a fluctuating quality. Such an issue has been reported in many state-of-the-art papers.

Diagnostic. For authoritative bodies, it is highly important to quantify the reliability of the information received. Such a need increases today with the number and range of data producers. Qualification consists in verifying whether the LC DBs fit with producers' and/or users' specifications in terms of semantic accuracy, geometric accuracy and completeness. This is all the more mandatory than it is widely recognized that some classes
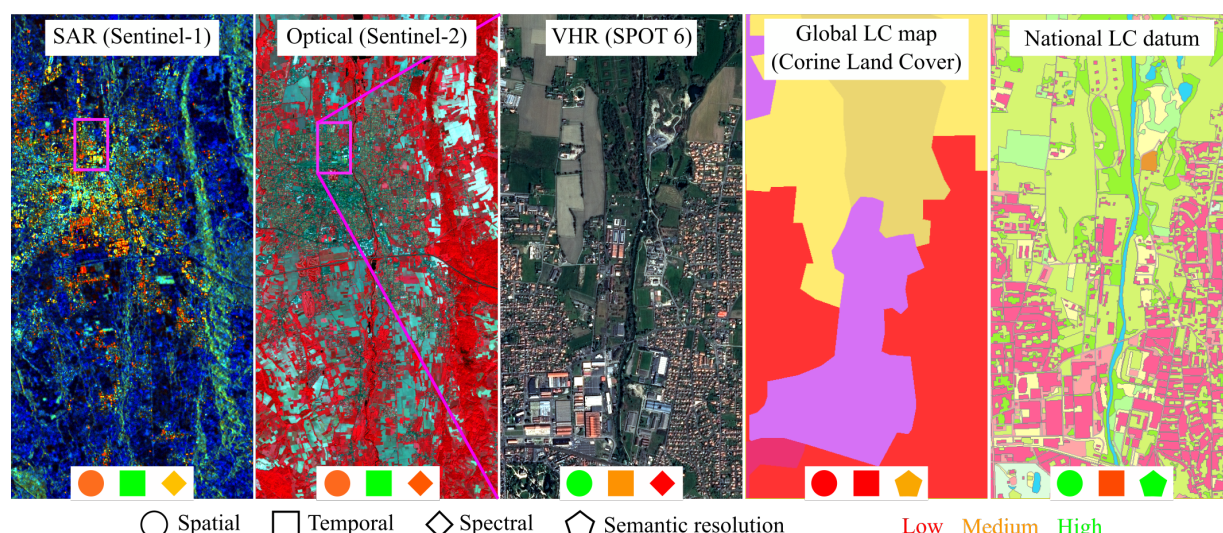
Figure 2. Main complementary data sources today for VHR land-cover mapping. The aim is to generate maps similar to a national LC datum (right) using (i) multi-modal images (time series and monoscopic VHR data) and (ii) existing imperfect geodatabases at multiple scales for the learning stage.

are more difficult to discriminate that the other ones at a given spatial and spectral resolution. In addition, nowadays, some LC DBs stem from the fusion of various disjoint automatic processes, sometimes for several contributors (Vargas-Munoz et al., 2019): they should be inspected before being broadcast. Finding reliable metrics is eventually not always straightforward.

Automation. Still today, many DBs are created with manual inputs: rules for fusing basic multi-source LC DBs, landscape specific parameter tuning, supervision in the discrimination of the classes of interest, manual improvements at the end of the process, etc. For efficient updating, upscaling and diagnostic, such an intervention should be minimized as much as possible. However, human input should not be totally discarded since it offers some invaluable advantages (Wuttke et al., 2018).

Harmonization. It first corresponds to a clear need to improve LC DB quality. Indeed, significant variations in class accuracy exist. The relative importance of different class accuracy also varies with the users. Secondly, there are significant differences in the LC DB generation approaches and, most of the time, poor agreement among datasets (i) at the same level and (ii) between spatial scales. It has been realized that these data products are difficult to match in terms of generation and semantics. Harmonization and interoperability initiatives are currently on-going, such as EAGLE (Kleeschulte et al., 2016).

2.3 Operational constraints

Operational mapping is opposed to *experimental* mapping. Experimental mapping refers to the development and assessment of new methods whereas operational techniques focus on process upscaling and reliable product delivery within a predefined time schedule (Chen et al., 2015). In terms of mapping issues for authoritative bodies, it first consists in automating/improving existing processes. Secondly, it targets to solve issues that can not be achieved in a reasonable cost or time, otherwise (field surveys and/or visual analysis). Consequently, operational effectiveness encompasses several remote sensing based challenges. Research for operational VHR land-cover product generation tackles more issues than for global mapping. This is due to two main facts: (i) a higher spatial accuracy

is required; (ii) LC products have to meet needs from various end-users and at several scales (national → local).

Automatic data processing. For versatility and scalability purposes, a very restricted number of parameters should exist for the developed methods. Optimally, they should be tuned automatically per area of interest. If impossible, it should be circumvented by a limited set of training areas.

Processing chain optimization. Each element of the land-cover generation workflow should be improved as much as possible, so as to increase discrimination accuracy with decreasing computing times. Meanwhile, the pipeline should be reproducible: they should remain as simple as possible in order to ease the transfer to the production services, i.e., units responsible for (i) the efficient implementation of the tools using the most adapted libraries and (ii) software running/data processing.

Versatility. Both for targeted and general LC mapping, methods should be adapted to various environments or at least should be easily transferable (genericity of the method). Additionally, they have to be locally relevant and at the same time globally consistent in order to achieve an accurate description of the national landscapes (genericity of the final product).

Upscaling. Country-wide classifications are now required for VHR maps in order to become a reliable substitute to topographic databases and current workflows based on visual inspection of images. A specific focus should be made on the best trade-off between acceptable processing times/simplicity, and satisfactory spatial and semantic accuracy.

Selection/evaluation of the remote sensing data sources. The first assessment challenge is related to the availability of remote sensing data at a regular basis through existing archives or dedicated Spatial Data Infrastructures. Many national mapping and cadastral agencies were used to process a specific kind of data with a given time lapse between two acquisitions. New datasets with increasing temporal and spectral resolutions can be beneficial. Secondly, experimental sensors may also provide observations of different kinds with ad hoc surveys (e.g., oblique imagery, hyperspectral images, or full-waveform lidar signals).

Their relevance should be evaluated in order to assess whether they could be integrated in the existing workflows.

Optimal exploitation of existing LC data. Existing land-cover geodatabases are an invaluable input even if (i) they may not be up-to-date or accurate enough and (ii) the existing nomenclature may not fit with the current one (Gressin et al., 2014). Nevertheless, they should be exploited since they may offer a reliable replacement solution to human operators as reference data and for parameter tuning and learning tasks.

Duality. LC databases can be created for mapping and statistical purposes (e.g., spatial indicators, or forest statistical inventory), that are not easily compatible: exhaustivity and representativeness are privileged, or conversely spatial accuracy. In a dual perspective, operational effectiveness should be able to deal with both objectives with few changes in the methodology. Many public policies, in particular related to urban areas (imperviousness), require both evolution metrics and accurate location of the underlying phenomena (Costa et al., 2018).

3. PARADIGM CHANGE IN LAND-COVER MAPPING

The context in remote sensing has significantly evolved over the five last years. We have witnessed a paradigm change that has raised new issues, shifting research to new processing domains.

Data is more and more easily available. In addition to continuously decreasing satellite data cost, images become more often free for non-commercial purposes. Open access data is advocated by many researchers, research institutes and public bodies. Large amounts of optical and radar images as well as 3D point clouds can be downloaded through dedicated Spatial Data Infrastructures and Web portals. Raw data being useless for many practitioners, existing platforms (Theia, France; EODC, Austria) are even augmented by products and services (e.g., bio-geophysical parameters).

In parallel, **open-access data has arrived** (Wulder, Coops, 2015). Satellite sensor archives are released for free, giving access to 30-40 years of collection of Earth Imagery (Landsat, SPOT (Ultré-Guerard, Boissin, 2015, Wulder et al., 2019)). Initiatives at local scales exist, allowing to tackle issues for specific environments and to envisage knowledge transfer. New sensors also provide data for free, such as the Sentinel program for the European Space Agency (ESA). The other significant advantage is the global coverage of such datasets, allowing to collect observations almost wherever around the globe. Many national mapping agencies also release their own data, giving access to older images with higher spatial resolutions. Free and open data policy is expected to foster data reuse (Zhu, Woodcock, 2014), and the development of new services in Earth Observation.

The variety of passive and active sensors proliferates in terms of spatial, spectral, temporal resolutions (Toth, Jóźków, 2016): hyperspectral sensors, time-series of images, video sequences from small satellites or Unmanned Aerial Systems, full-waveform and multispectral lidar, radar with full polarimetric capability. The larger agility of spaceborne optical sensors (Pléiades, MISR, CHRIS, Worldview-2/3) allows to give access to multi-view multi-angular datasets. Digital cameras for airborne acquisitions are getting finer spatial resolutions with an increasing number of spectral bands. Oblique imagery is more widespread. Technological evolution has been driven by experience collected after first decades of remote sensing and geodatabase use. Subsequently, remote sensing is now inherently multi-modal (Gomez-Chova et al., 2015): complementary observations can be exploited and can mitigate limitations of a particular sensor (Joshi et al., 2016). This is the concept of virtual constellations (Wulder et al., 2019).

Many national mapping agencies and local public bodies release reference topographic, land-cover and land-use databases for free at many levels (e.g., Globeland30, ESA Copernicus program, US NLCD). The need for updating, diagnostic, and refinement has emerged since initial DBs are now established, leading to clearer requirements (see Section 2.2).

An increasing number of free and open-source processing softwares for image and 3D point cloud exploitation becomes available (Christophe, Inglada, 2009). This allows to give access to core processing techniques (calibration, registration, classification, segmentation). It permits to build higher level processing chains that can be easily evaluated with the growing availability of contest and benchmark datasets (Rottensteiner et al., 2014, Braun et al., 2015, Maggiori et al., 2017, Demir et al., 2018, Azimi et al., 2019, IEEE GRSS Data Fusion Contest, 2020). However, such initiatives are not sufficiently spread and are still of limited scope: sensor-specific, few efficient computational capacities, limited interfacing with high-level programming languages etc. The situation is steadily improving with the great availability of deep neural network (DNN) frameworks (Robinson et al., 2019a), within well-documented libraries (Zhu et al., 2017).

It is accompanied by the set up of multiple High Performance Computing infrastructures (HPC) and libraries, mandatory so as to upscale proposed workflows, swallow the amount of geospatial images, and provide maps in decent times (Chi et al., 2016, Haut et al., 2019). Infrastructures are located either in universities/institutes and national data centers, or are now directly provided by/rent to private companies (Google, Amazon Web Services, Microsoft Azure).

4. CURRENT CHALLENGES

The keywords are accuracy and scalability: reliable information layers, temporally and spatially homogeneous over a large territory, ideally country-scale and whatever the nomenclature. Main perspectives can be decomposed into: (i) New methods for improving classification and segmentation accuracy with upscaling ability; (ii) Optimal exploitation of multiple data sources (correlated with (i)), and (iii) fostering applications of scene interpretation from land-cover and land-use mapping.

4.1 Efficient classifiers

The two expected features of classifiers are accuracy and scalability. Higher accuracies can be obtained with automatic or manual knowledge integration as well as a better use of existing land-cover and remote sensing datasets (next section).

A better training stage? All workflows are based on the supervised classification of remote sensing image(s). Therefore, they heavily rely on existing reference data (Section 4.2) and on efficient training procedures. Most fail today in dealing with imbalanced datasets and rare classes: at the data level, standard under/oversampling often lead to over-fitting or redundancy. Cost-sensitive learning is better suited. Continuously learning class costs leads to intractable optimization with large scale

Criterion (↓) / Scenario (→)	Experimental mapping	Operational mapping		
		HR	VHR	VHR today
Spatial accuracy	🧩 - 🔒	🧩 - 🔒	🧩 - 🔒	🧩 - 🔒
Semantic accuracy	🧩 - 🔒	🧩 - 🔒	🧩 - 🔒	🧩 - 🔒
Updateness & change detection	🧩 - 🔒	🧩 - 🔒	🧩 - 🔒	🧩 - 🔒
Automation	🧩 - 🔒	🧩 - 🔒	🧩 - 🔒	🧩 - 🔒
Simplicity	🧩 - 🔒	🧩 - 🔒	🧩 - 🔒	🧩 - 🔒
Versatility	🧩 - 🔒	🧩 - 🔒	🧩 - 🔒	🧩 - 🔒
Upscaling	🧩 - 🔒	🧩 - 🔒	🧩 - 🔒	🧩 - 🔒
RS data evaluation/fusion	🧩 - 🔒	🧩 - 🔒	🧩 - 🔒	🧩 - 🔒
Exploitation of existing databases	🧩 - 🔒	🧩 - 🔒	🧩 - 🔒	🧩 - 🔒

Table 1. Main challenges in land-cover mapping. For operational cases, we distinguish global/High Resolution (HR), national/Very High Resolution (VHR), and VHR with current paradigm change (VHR today). 🧩 and 🔒 indicate the importance and the complexity of each criterion, respectively. The color quantifies each of them (low - medium - high).

datasets (Huang et al., 2016). Authors prefer defining ad hoc losses, e.g., with hard negative mining (Dong et al., 2017), if possible supported by joint clustering (Hayat et al., 2019). Similarly, unseen classes are not handled when applying existing classifiers on new areas, preventing their upscaling ability. *Few-shot learning* and *weakly supervised* solutions (Gidaris, Komodakis, 2019) can cope with such issues, handle tail classes (Liu et al., 2019), can help in domain adaptation tasks, and in reducing the required amount of training in DNN, in particular when integrating unlabelled yet useful data. Curriculum learning may also be adopted so as to sequentially handle classification tasks (Wang et al., 2019). Such strategies have been barely addressed in the remote sensing literature (Kellenberger et al., 2019). In parallel, *Incremental* and *Meta-learning* provide suitable solutions for adding new classes and reference data (Wang et al., 2017, Tasar et al., 2019). Catastrophic forgetting can be avoided while it remains complex to generalize to new areas.

Deep-learning architectures. Recent DNN approaches have shown great performance in processing either VHR optical images (for anthropic classes) or time-series of HR images (for natural classes: forests and crops), alleviating the tedious design of hand-crafted features. Data sources and architectures exhibit complementary strengths that should be merged in order to provide a genuine multi-scale, multi-resolution multi-modal multi-class framework. First attempts in this direction have been proposed (Benedetti et al., 2018) and there is still room for improvement in order to propose (i) fully agnostic pipelines (Perez-Rua et al., 2019), (ii) that can be fed with other modalities (e.g., 3D point clouds). In addition, more complex architectures often come with larger networks and an increasing number of parameters. Even if outstanding computational resources are now available, for simplicity and scalability, it remains relevant to propose simpler models and compressed solution. *Distillation* is a suitable solution in such a direction (Hinton et al., 2015, Chen et al., 2018, Liu et al., 2019).

Vector semantic segmentation. Regularization solutions are adopted in the literature to cope with noisy pixel-wise classifications and fuzzy Convolutional Neural Networks (CNN) predictions. This remains unsatisfactory in VHR mapping since sharp contours are not retrieved. This leads to tedious ad hoc post-processing, reducing both automation and versatility. The classification process is commonly eased either with segmentation algorithms that provide strong local spatial supports, sometimes at various scales or directly with *semantic segmentation*. This aims to solve the interleaved issue of classification and segmentation by combining top-down and bottom-up cues (Derksen et al., 2018, Marmanis et al., 2018). CNN is now a golden standard. Remaining challenges revolve around two main goals. First, raster maps provided by are not decent mapping objects since vector data and instances are required. Current efforts are put towards generating directly vector maps with individual polygonal objects from CNN architectures (Girard, Tarabalka, 2018, Li et al., 2019) or after a coarse vectorization step (Li et al., 2020b). Topology can be preserved. So far, no framework was able to derive a full partition of the space, over large scales and multiple classes. Secondly, label prediction can now be accompanied with other tasks that can be fruitful for mapping purposes. In deep-based architectures, one can now couple networks that share computation and knowledge for the benefit of several tasks: *multi-task learning*. It is beneficial for instance for unseen image generation or data completion such as Digital Surface Model (Carvalho et al., 2019), and should now be extended to other modalities. A particular case, and a major trend in computer vision, is *panoptic segmentation* (Kirillov et al., 2019), which includes instance segmentation. It should become prominent in land cover mapping: both statistical (object counting) and delineation (dense labelling) requirements can be met. Efficient solutions such as Mask R-CNN can definitively help (Hu et al., 2018).

4.2 Optimal exploitation of data sources

Data fusion Efficient multi-class semantic segmentation for land-cover mapping requires a synergistic use of all available remote sensing data. Higher semantic accuracy is now possible with HR multitemporal sensors but to the detriment of spatial accuracy. Both mid- and late-fusion solutions are conceivable: (i) an agnostic reasoning that requires novel deep-based architectures (see above), or (ii) an educated hierarchical strategy: each family of images can be processed separately with the optimal classifier, the remaining challenge being finding the most suitable solution for late fusion. Again, multi-task learning can be envisaged as well as standard fusion techniques or more theoretically Multi-Armed Bandits (Radlinski et al., 2008). Note that VHR data may not be available in all areas every year. This requires DNN able to handle missing modalities (Kampffmeyer et al., 2018).

Alternatively, low-level fusion is insufficiently addressed. The *data cube* paradigm (Augustin et al., 2019) in the multi-modal case is limited to the *stack-and-classify* approach, which consists in resampling all data to the highest resolution, either at ingestion time or at query time. In case of very distinct modalities, super-resolution or disaggregation techniques based on the underlying physics are preferred, pan-sharpening solutions being the most widespread. However, in order to minimize the loss of information and to propose an interpretable solution (required by many end-users, in particular for classifica-

tion/change detection assessment), both novel physical and statistical solutions have to be proposed. Today, variational autoencoders appear to be a suitable solution to find such an optimal common representation space (Sønderby et al., 2016).

Humans in the loop. Crowdsourcing has already proved to be efficient for label collection and classification refinement. *Off-line* solutions ignore the classification task; *on-line* methods would leverage the impact of the annotation effort and improve classification performances (Cui et al., 2016). The initial collection of training data can be reduced (Robinson et al., 2019b), one can adapt the classification task to particular DB specifications (see below), and interactions with the classifier would help collecting samples in areas prone to errors, favoring *incremental learning*. Active learning strategies are not new (Laroze et al., 2018). However, they have never been deployed under operational constraints, where focus should be put with reduced prediction time and uncertainty analysis so as propose a fast yet efficient Human Computer Interaction framework.

Exploitation of existing geospatial databases These are invaluable sources for training classifiers. They are both available at large scales with limited nomenclatures or over limited areas with fine-grained categorization. In the latter case, if one aims to adopt such set of classes, it becomes a *weakly-supervised* or a *transfer learning* issue. Most of the literature focuses in transferring existing knowledge to new sensors; additional effort should be put on the complementary task (Redko et al., 2019), especially foreseeing the cases when reference data will no longer be available (Tardy et al., 2019). Furthermore, most LC DBs are organized in a hierarchical way (forests coniferous/deciduous/other). Such taxonomy has not been yet exploited in order to constraint the classification problem (Verma et al., 2012), or, in an pragmatic reasoning, to find the most adapted level of representation (label set) given a RS image. Recent developments show high potential (Chami et al., 2019).

4.3 Fostering applications

Towards land-use mapping. This remains a remarkably ignored domain while it has been shown to be the most consuming and challenging task in geodatabase generation. Many reference data exist, allowing a suitable training step. The main bottleneck is not a correct local discrimination but an efficient partitioning of the space, especially in urban areas. This is again a *semantic segmentation* task. Finding borders between land-use classes is not trivial, which explains why researchers often complement their discriminative workflow with ground-based images (Srivastava et al., 2019).

Land-cover dynamics monitoring. Mono-date classification is now longer sufficient and should be accompanied with change detection (Wulder et al., 2018). For robust multi-date comparisons, the challenges consist first in integrating into the (structured) classifier the knowledge about the conceivable trajectory of changes (Bailly et al., 2018). Secondly, efficient methods should adequately handle LC DB specifications in order to avoid producing a large number of false alarms (Gressin et al., 2014). Such an issue has not been yet included in the current solutions. Eventually, near real-time change detection are often desired by stakeholders for a larger number of classes with high stakes (forests, crops), which has been made possible with the temporal resolution of Landsat and Sentinel (Zhu, Woodcock, 2014). Solutions exist for simple configurations (Dutrieux et al., 2015). Under operational constraints, one first has to generate a time-series of cloudless observations and to develop a weakly supervised or unsupervised breakpoint detection method, so as to cope with missing or very limited reference data (Griffiths et al., 2020, Li et al., 2020a).

5. CONCLUSIONS

In this paper, we presented the current status of Very High spatial Resolution land-cover mapping under operational constraints. We discussed and quantified main features and requirements and introduced current challenges. We advocate for increasing efforts in upscaling current methodological solutions and in designing workflows integrating human knowledge.

ACKNOWLEDGMENTS

REFERENCES

Aschbacher, J., Milagro-Pérez, M. P., 2012. The European Earth monitoring (GMES) programme: Status and perspectives. *Remote Sensing of Environment*, 120, 3-8.

Augustin, H., Sudmanns, M., Tiede, D., Lang, S., Baraldi, A., 2019. Semantic Earth Observation Data Cubes. *Data*, 4(3).

Azimi, S. M., Henry, C., Sommer, L., Schumann, A., Vig, E., 2019. Skyscapes fine-grained semantic understanding of aerial scenes. *ICCV*.

Bailly, S., Giordano, S., Landrieu, L., Chehata, N., 2018. Crop-rotation structured classification using multi-source Sentinel images and LPIS for crop type mapping. *IGARSS*, 1950–1953.

Ban, Y., Gong, P., Giri, C., 2015. Global land cover mapping using Earth observation satellite data: Recent progresses and challenges. *ISPRS Journal of Photogrammetry and Remote Sensing*, 103, 1-6.

Benedetti, P., Ienco, D., Gaetano, R., Ose, K., Pensa, R. G., Dupuy, S., 2018. $M^3 Fusion$: A Deep Learning Architecture for Multiscale Multimodal Multitemporal Satellite Data Fusion. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 11(12), 4939-4949.

Bontemps, S., Herold, M., Kooistra, L., van Groenestijn, A., Hartley, A., Arino, O., Moreau, I., Defourny, P., 2012. Revisiting land cover observation to address the needs of the climate modeling community. *Biogeosciences*, 9(6), 2145-2157.

Braun, A. C., Weinmann, M., Keller, S., Müller, R., Reinartz, P., Hinz, S., 2015. The EnMAP contest: Developing and comparing classification approaches for the environmental mapping and analysis programme – dataset and first results. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XL-3/W3, 169-175.

Brovelli, M. et al., 2018. A crowdsourcing-based game for land cover validation. *Applied Geomatics*, 10, 1-11.

Buchanan, G., Nelson, A., Mayaux, P., Hartley, A., Donald, P., 2009. Delivering a Global, Terrestrial, Biodiversity Observation System through Remote Sensing. *Conservation Biology*, 23(2), 499-502.

Carvalho, M., Le Saux, B., Trouvé-Peloux, P., Champagnat, F., Almansa, A., 2019. Multitask Learning of Height and Semantics From Aerial Images. *IEEE Geoscience and Remote Sensing Letters*, 1-5.

Chami, I., Ying, R., Ré, C., Leskovec, J., 2019. Hyperbolic graph convolutional neural networks. *NeurIPS*.

Chen, G., Zhang, X., Tan, X., Cheng, Y., Dai, F., Zhu, K., Gong, Y., Wang, Q., 2018. Training Small Networks for Scene Classification of Remote Sensing Images via Knowledge Distillation. *Remote Sensing*, 10(5), 719.

Chen, J., Chen, J., Liao, A., Cao, X., Chen, L., Chen, X., He, C., Han, G., Peng, S., Lu, M., Zhang, W., Tong, X., Mills, J., 2015. Global land cover mapping at 30 m resolution: A POK-based operational approach. *ISPRS Journal of Photogrammetry and Remote Sensing*, 103, 7-27.

Chi, M., Plaza, A., Benediktsson, J., Sun, Z., Shen, J., Zhu, Y., 2016. Big Data for Remote Sensing: Challenges and Opportunities. *Proceedings of the IEEE*, 104(11), 2207-2219.

Christophe, E., Inglada, J., 2009. Open Source Remote Sensing: Increasing the Usability of Cutting-Edge Algorithms. *IEEE Geoscience and Remote Sensing Newsletter*, 35(5), 9-15.

Congalton, R., Gu, J., Yadav, K., Thenkabail, P., Ozdogan, M., 2014. Global Land Cover Mapping: A Review and Uncertainty Analysis. *Remote Sensing*, 6(12), 12070-12093.

Costa, H., Almeida, D., Vala, F., Marcelino, F., Caetano, M., 2018. Land Cover Mapping from Remotely Sensed and Auxiliary Data for Harmonized Official Statistics. *ISPRS International Journal of Geo-Information*, 7(4).

Cui, Y., Zhou, F., Lin, Y., Belongie, S., 2016. Fine-grained categorization and dataset bootstrapping using deep metric learning with humans in the loop. *CVPR*.

Demir, I., Koperski, K., Lindenbaum, D., Pang, G., Huang, J., Basu, S., Hughes, F., Tuia, D., Raska, R., 2018. Deepglobe 2018: A challenge to parse the Earth through satellite images. *CVPR Workshops*.

Derksen, D., Inglada, J., Michel, J., 2018. Spatially precise contextual features based on superpixel neighborhoods for land cover mapping with high resolution satellite image time series. *IGARSS*, 200–203.

Dong, Q., Gong, S., Zhu, X., 2017. Class rectification hard mining for imbalanced deep learning. *ICCV*.

Dutrieux, L., Verbesselt, J., Kooistra, L., Herold, M., 2015. Monitoring forest cover loss using multiple data streams, a case study of a tropical dry forest in Bolivia. *ISPRS Journal of Photogrammetry and Remote Sensing*, 107, 112-125.

Feddema, J., Oleson, K., Bonan, G., Mearns, L., Buja, L., Meehl, G., Washington, W., 2005. The Importance of Land-Cover Change in Simulating Future Climates. *Science*, 310, 1674-1678.

Gidaris, S., Komodakis, N., 2019. Generating classification weights with GNN denoising autoencoders for few-shot learning. *CVPR*.

Girard, N., Tarabalka, Y., 2018. End-to-end learning of polygons for remote sensing image classification. *IGARSS*, 2083–2086.

Giri, C., 2012. *Remote sensing of land use and land cover: Principles and applications*. CRC Press, Boca Raton, FL, USA, chapter Brief Overview of Remote Sensing of Land Cover.

Giri, C., Pengra, B., Long, J., Loveland, T., 2013. Next generation of global land cover characterization, mapping, and monitoring. *International Journal of Applied Earth Observation and Geoinformation*, 25, 30-37.

Gomez-Chova, L., Tuia, D., Moser, G., Camps-Valls, G., 2015. Multimodal Classification of Remote Sensing Images: A Review and Future Directions. *Proceedings of the IEEE*, 103(9), 1560-1584.

Grekousis, G., Mountrakis, G., Kavouras, M., 2015. An overview of 21 global and 43 regional land-cover mapping products. *International Journal of Remote Sensing*, 36(21), 5309-5335.

Gressin, A., Mallet, C., Vincent, N., Paparoditis, N., 2014. Updating the new French national land cover database. *IGARSS*, 3534–3537.

Griffiths, P., Nendel, C., Pickert, J., Hostert, P., 2020. Towards national-scale characterization of grassland use intensity from integrated Sentinel-2 and Landsat time series. *Remote Sensing of Environment*, 238, 111-124.

Haut, J. M., Gallardo, J. A., Paoletti, M. E., Cavallaro, G., Plaza, J., Plaza, A., Riedel, M., 2019. Cloud Deep Networks for Hyperspectral Image Analysis. *IEEE Transactions on Geoscience and Remote Sensing*, 57(12), 9832-9848.

Hayat, M., Khan, S., Zamir, S. W., Shen, J., Shao, L., 2019. Gaussian affinity for max-margin class imbalanced learning. *ICCV*.

Hinton, G., Vinyals, O., Dean, J., 2015. Distilling the knowledge in a neural network. *NIPS Deep Learning and Representation Learning Workshop*.

Homer, C., Dewitz, J., Yang, L., Jin, S., Danielson, P., Xian, G., Coulston, J., Herold, N., Wickham, J., Megown, K., 2015. Completion of the 2011 National Land Cover Database for the conterminous United States - Representing a decade of land cover change information. *Photogrammetric Engineering & Remote Sensing*, 81(5), 345-354.

Hu, R., Dollár, P., He, K., Darrell, T., Girshick, R., 2018. Learning to segment every thing. *CVPR*.

Huang, C., Li, Y., Change Loy, C., Tang, X., 2016. Learning deep representation for imbalanced classification. *CVPR*.

IEEE GRSS Data Fusion Contest, 2020. http://www.grss-ieee.org/community/technical-committees/data-fusion/data-fusion-contest/. Access Feb. 3, 2020.

Inglada, J., Vincent, A., Arias, M., Tardy, B., Morin, D., Rodes, I., 2017. Operational High Resolution Land Cover Map Production at the Country Scale Using Satellite Image Time Series. *Remote Sensing*, 9(1).

Jalal, R. et al., 2019. Toward Efficient Land Cover Mapping: An Overview of the National Land Representation System and Land Cover Map 2015 of Bangladesh. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 12(10), 3852-3861.

Joshi, N., Baumann, M., Ehammer, A., Fensholt, R., Grogan, K., Hostert, P., Jepsen, M. R., Kuemmerle, T., Meyfroidt, P., Mitchard, E. T. A., Reiche, J., Ryan, C. M., Waske, B., 2016. A Review of the Application of Optical and Radar Remote Sensing Data Fusion to Land Use Mapping and Monitoring. *Remote Sensing*, 8(1), 70.

Kampffmeyer, M., Salberg, A., Jenssen, R., 2018. Urban Land Cover Classification With Missing Data Modalities Using Deep Convolutional Neural Networks. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 11(6), 1758-1768.

Kellenberger, B., Marcos, D., Lobry, S., Tuia, D., 2019. Half a Percent of Labels is Enough: Efficient Animal Detection in UAV Imagery Using Deep CNNs and Active Learning. *IEEE Transactions on Geoscience and Remote Sensing*, 57(12), 9524–9533.

Kirillov, A., Girshick, R., He, K., Dollar, P., 2019. Panoptic feature pyramid networks. *CVPR*.

Kleeschulte, S., Soukup, T., Hazeu, G., Smith, G., Arnold, S., Kosztra, B., 2016. *The future of land monitoring in Europe*. CRC Press.

Laroze, M., Dambreville, R., Friguet, C., Kijak, E., Lefèvre, S., 2018. Active learning to assist annotation of aerial images in environmental surveys. *CBMI*.

Li, J., Huang, X., Chang, X., 2020a. A label-noise robust active learning sample collection method for multi-temporal urban land-cover classification and change analysis. *ISPRS Journal of Photogrammetry and Remote Sensing*, 163, 1 - 17.

Li, M., Lafarge, F., Marlet, R., 2020b. Approximating shapes in images with low-complexity polygons. *CVPR*.

Li, Z., Wegner, J. D., Lucchi, A., 2019. Topological map extraction from overhead images. *ICCV*.

Liu, Y., Chen, K., Liu, C., Qin, Z., Luo, Z., Wang, J., 2019. Structured knowledge distillation for semantic segmentation. *CVPR*.

Maggiori, E., Tarabalka, Y., Charpiat, G., Alliez, P., 2017. Can semantic labeling methods generalize to any city? the INRIA aerial image labeling benchmark. *IGARSS*.

Manakos, I., Braun, M. (eds), 2014. *Land Use and Land Cover Mapping in Europe: Practices & Trends*. Springer.

Marmanis, D., Schindler, K., Wegner, J., Galliani, S., Datcu, M., Stilla, U., 2018. Classification with an edge: Improving semantic image segmentation with boundary detection. *ISPRS Journal of Photogrammetry and Remote Sensing*, 135, 158-172.

Pereira, H. et al., 2013. Essential Biodiversity Variables. *Science*, 339(6117), 277-278.

Perez-Rua, J.-M., Vielzeuf, V., Pateux, S., Baccouche, M., Jurie, F., 2019. MFAS: Multimodal Fusion Architecture Search. *CVPR*.

Radlinski, F., Kleinberg, R., Joachims, T., 2008. Learning diverse rankings with multi-armed bandits. *ICML*.

Redko, I., Courty, N., Flamary, R., Tuia, D., 2019. Optimal Transport for Multi-source Domain Adaptation under Target Shift. *AISTATS*, 89.

Robinson, C., Hou, L., Malkin, K., Soobitsky, R., Czawlytko, J., Dilkina, B., Jojic, N., 2019a. Large scale high-resolution land cover mapping with multi-resolution data. *CVPR*.

Robinson, C., Ortiz, A., Malkin, K., Elias, B., Peng, A., Morris, D., Dilkina, B. N., Jojic, N., 2019b. Human-machine collaboration for fast land cover mapping. *AAAI*.

Rottensteiner, F., Sohn, G., Gerke, M., Wegner, J. D., Breitkopf, U., Jung, J., 2014. Results of the ISPRS benchmark on urban object detection and 3D building reconstruction. *ISPRS Journal of Photogrammetry and Remote Sensing*, 93, 256-271.

See, L., Perger, C., Hofer, M., Weichselbaum, J., Dresel, C., Fritz, S., 2015. LACO-Wiki: An open-access online portal for land-cover validation. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, II-3/W5, 167-171.

Skidmore, A., Pettorelli, N., Coops, N., Geller, G., Hansen, M., Lucas, R., Mücher, C., O'Connor, B., Paganini, M., Pereira, H., Schaepman, M., Turner, W., Wang, T., Wegmann, M., 2015. Environmental science: Agree on biodiversity metrics to track from space. *Nature*, 523(7561), 30-31.

Sønderby, C. K., Raiko, T., Maaløe, L., Sønderby, S. K., Winther, O., 2016. Ladder variational autoencoders. *NIPS*.

Srivastava, S., Vargas-Muñoz, J. E., Tuia, D., 2019. Understanding urban landuse from the above and ground perspectives: A deep learning, multimodal solution. *Remote Sensing of Environment*, 228, 129 - 143.

Tardy, B., Inglada, J., Michel, J., 2019. Assessment of Optimal Transport for Operational Land-Cover Mapping Using High-Resolution Satellite Images Time Series without Reference Data of the Mapping Period. *Remote Sensing*, 11(9), 1047.

Tasar, O., Tarabalka, Y., Alliez, P., 2019. Incremental Learning for Semantic Segmentation of Large-Scale Remote Sensing Data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 12(9), 3524-3537.

Tchuenté, A. K., Roujean, J.-L., Jong, S. D., 2011. Comparison and relative quality assessment of the GLC2000, GLOBCOVER, MODIS and ECOCLIMAP land cover data sets at the African continental scale. *International Journal of Applied Earth Observation and Geoinformation*, 13(2), 207-219.

Tewkesbury, A. P., Comber, A. J., Tate, N. J., Lamb, A., Fisher, P. F., 2015. A critical synthesis of remotely sensed optical image change detection techniques. *Remote Sensing of Environment*, 160, 1-14.

Toth, C., Jóźków, G., 2016. Remote sensing platforms and sensors: A survey. *ISPRS Journal of Photogrammetry and Remote Sensing*, 115, 22-36.

Tsendbazar, N., de Bruin, S., Mora, B., Schouten, L., Herold, M., 2016. Comparative assessment of thematic accuracy of GLC maps for specific applications using existing reference data. *International Journal of Applied Earth Observation and Geoinformation*, 44, 124-135.

Ultré-Guerard, P., Boissin, B., 2015. The CNES Earth Observation Program. *IEEE Geoscience and Remote Sensing Magazine*, 3(2), 41-50.

Vargas-Munoz, J. E., Lobry, S., Falcao, A. X., Tuia, D., 2019. Correcting rural building annotations in OpenStreetMap using convolutional neural networks. *ISPRS Journal of Photogrammetry and Remote Sensing*, 147, 283 - 293.

Verma, N., Mahajan, D., Sellamanickam, S., Nair, V., 2012. Learning hierarchical similarity metrics. *CVPR*.

Wang, Y., Gan, W., Yang, J., Wu, W., Yan, J., 2019. Dynamic curriculum learning for imbalanced data classification. *ICCV cy*.

Wang, Y.-X., Ramanan, D., Hebert, M., 2017. Learning to model the tail. *NIPS*.

Wulder, M. A., Coops, N. C., 2015. Satellites: Make Earth observations open access. *Nature*, 513(7516), 30-31.

Wulder, M. A. et al., 2019. Current status of Landsat program, science, and applications. *Remote Sensing of Environment*, 225, 127 - 147.

Wulder, M., Coops, N., Roy, D., White, J., Hermosilla, T., 2018. Land cover 2.0. *International Journal of Remote Sensing*, 39(12), 4254-4284.

Wuttke, S., Middelmann, W., Stilla, U., 2018. Improving the Efficiency of Land Cover Classification by Combining Segmentation, Hierarchical Clustering, and Active Learning. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 11(11), 4016-4031.

Yang, L. et al., 2018. A new generation of the United States National Land Cover Database: Requirements, research priorities, design, and implementation strategies. *ISPRS Journal of Photogrammetry and Remote Sensing*, 146, 108 - 123.

Yu, L. et al., 2014. Meta-discoveries from a synthesis of satellite-based land-cover mapping research. *International Journal of Remote Sensing*, 35(13), 4573-4588.

Zhu, X. X., Tuia, D., Mou, L., Xia, G., Zhang, L., Xu, F., Fraundorfer, F., 2017. Deep Learning in Remote Sensing: A Comprehensive Review and List of Resources. *IEEE Geoscience and Remote Sensing Magazine*, 5(4), 8-36.

Zhu, Z., Woodcock, C., 2014. Continuous change detection and classification of land cover using all available Landsat data. *Remote Sensing of Environment*, 144, 152-171.