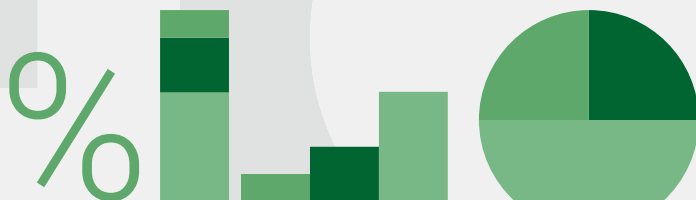




FSO News



02 Territory and environment

Neuchâtel, september 2023

The use of artificial intelligence in the FSO's land use statistics

The ADELE deep learning tool: how it works, capabilities and limitations

Since 2022, artificial intelligence has been contributing for the first time "in real use" to the production of an official statistic in Switzerland. The deep learning tool ADELE has enabled partial automation of the labour-intensive task of interpreting the aerial photographs on which Switzerland's land use statistics are based. ADELE is an excellent example of the new possibilities offered by AI in the production of official statistics and of how AI could be used in practice. However, ADELE has also highlighted certain conflicts between the way artificial intelligence works and the need to produce reliable series of statistics for use in the long term.

1 The FSO's land use statistics: Content and purpose

The land use statistics produced by the Federal Statistical Office (FSO) provide information about land use and land cover in Switzerland and on how these change over time. It is therefore an essential tool for long-term spatial monitoring. The findings of the land use statistics make it possible to assess whether and to what extent the changes in land use coincide with the objectives of Swiss spatial development policy and hence also with the desire for a more economic approach to land as a resource.

Currently, the results of four methodologically consistent and completed surveys are available: 1985, 1997, 2009 and 2018 (see also box on right). Each date refers to the final year of a multi-year survey period. Another survey began in 2021, the first results of which (cantons of Geneva and Vaud) were published in summer 2023. Results for the whole of Switzerland are expected to be available in 2027.

Land use statistics: findings and detailed methodology documents

The graphic below shows an example of an analysis based on the data from the land use statistics survey. Between 1985 and 2018, an area of 0.7 square metres was lost every second to building development, meaning it was converted into land for settlement and urban areas.

Visit the FSO website at: www.landuse-stat.admin.ch for more land use graphics, detailed tables, explanations and methodology documents.

Change in land use in m² per second between 1985 and 2018

Increase and decrease by main category



Source: FSO – land use statistics (AREA)

© FSO 2023

2 Data sources and traditional survey process

Based on aerial photos

The basis for the land use statistics are high-resolution aerial photographs, which cover the whole of Switzerland's territory and which the FSO receives from the Federal Office of Topography. The images are taken by specially equipped aircraft that fly over Switzerland at regular intervals. The photographs are mainly used for creating and updating Swiss maps. The production of the land use statistics is one of many secondary uses.

Over 4 million sample points

The aerial photographs are first overlaid with a grid of sample points spaced 100 meters apart. This provides some 4.1 million sample points on which to base the statistics. The points are georeferenced and remain fixed. They are thus in the same position in every survey. In each survey, land use and land cover are determined anew for each individual sample point. The findings are then extrapolated to the whole territory.

Visual interpretation of the aerial photographs

The type of land use and land cover at the sample points is traditionally interpreted by humans. Trained analysts view the aerial photographs on a special stereoscopic 3D monitor and label each sample point with a predefined land use and land cover classification. The analysts examine the images of the area immediately surrounding a sample point. Is the sample point on a building? If so, what kind of building? Is it a residential, industrial,

or agricultural building? They distinguish between 46 different types of land use and 27 types of land cover (e.g. "buildings" or "shrubs"). The different types of land use all fall into one of four main categories: settlement and urban areas, agricultural areas, wooded areas and unproductive areas.

Additional information for interpreting the images

To make it easier for the analysts to determine the land use and cover at the sample point, they can draw on various additional information displayed on a second screen. The Federal Register of Buildings and Dwellings, for example, enables distinction between single-family and multi-family dwellings. The Topographic Landscape Model of Switzerland provides information on road width, and the Inventory of Utilised Agricultural Land helps identify special crops such as orchards or vineyards. Although the analysts have access to this information and a lot of other data, it is not always possible to pinpoint the exact use and land cover at a particular sample point. For such cases, an on-site inspection is needed.



An analyst interpreting the aerial photos on a stereoscopic 3D monitor. The two side screens display additional information.



An aerial image with sample points and codes indicating the determined land use. Code 106/11 stands for "consolidated surface in public building area".

3 The use of artificial intelligence in the FSO's land use statistics: potential and challenges

Despite the above-mentioned additional information and the use of certain aids (e.g. automatic plausibility check when entering the label information), the visual interpretation of the aerial photographs remains extremely laborious in view of the vast number of sample points – 4.1 million. It therefore makes sense to partially automate the task of interpretation using an image recognition program. Thanks to AI and machine learning (see below for details), automatic image recognition and the possibilities it offers have vastly improved in recent years.

Using AI in Switzerland's land use statistics appears to make even more sense if you consider the vast amount of existing material that can be used to train the models. Millions of aerial images have already been interpreted by humans over the course of past surveys and each individual interpretation and corresponding image has been stored digitally. The 'click work' that is necessary to train automated image recognition programs has thus already been done.

Considering this potential, the FSO commissioned a feasibility study in 2017 to clarify whether the interpretation of the aerial photographs that form the basis for Switzerland's land use statistics could be automatised at least partially using AI, aiming also to carry out the surveys at closer intervals in future. It quickly became apparent, however, that the potential is great but so are the challenges:

- Land use statistics require much higher precision than many commercial uses of image recognition.
- The huge number of sample points to interpret requires a similarly large amount of processing power and time. If it took one second to determine the land use and cover at a single point, it would take 47 days to determine 4.1 million survey points!
- Some land use and land cover categories look very similar on the aerial images, so additional information is needed to identify them. The AI model must therefore have the capacity to process other information besides images. This requires a more complex program architecture and even more processing power.
- The frequency with which different uses and cover categories occur varies considerably. Some categories occur so rarely that there is not enough material on them to train the AI with high precision. There is also a danger that these rare categories will be overlooked by the machine learning algorithms. Statistical methods must therefore be used to take into account the frequency with which the categories occur.

4 The ADELE AI tool: its structure and how it works in theory

Owing in part to the difficulties described above, it is unlikely that the interpretation of aerial images using AI will be fully automated in the short to medium term. Nevertheless, the FSO decided to develop a suitable AI application together with expert partners, as even partial automation would considerably speed up the task of interpreting the images. Because its core module uses the deep learning method, the AI tool was named ADELE, which stands for **"Arealstatistik DEep Learning"** (Arealstatistik = land use statistics).

Next, we will begin by describing how ADELE is able to automatically identify the land use and cover type in theory, i.e. regardless of the tool's practical limitations. We will then go on to show how ADELE is combined in practice with human interpretations of the images to speed up the process of generating the land use statistics without compromising the quality and reliability of the results.

4.1 ADELE's core: a deep learning module for the "visual" interpretation of aerial images

As we have seen, human interpreters rely primarily on the information contained in the aerial photographs to determine the land use and land cover at the sample points. A visual interpretation is carried out. The ADELE model also infers a land use and a land cover based on the images. Its core capability is to search for patterns in the area around the sample points on the same aerial photographs, allowing the points to be assigned to one of the 46 land use and 27 land cover categories of the land use statistics. For each sample point, it analyses a 50x50m section of image, which is equivalent to 200x200 (40,000) three-colour pixels. The different colour values form the basis for ADELE's interpretation of the image.

But how does ADELE interpret the images? The answer to this question is both simple and complicated. ADELE determines the land use and land cover by means of a statistical model consisting of a vast number of parameters set by the tool itself during the training phase. ADELE has not been fed any formal rules like "colour white = glacier" by human programmers. The reason for this is that it simply isn't possible to manually program all the rules needed for the correct interpretation of the aerial images. Instead, humans have programmed the tool to "train" itself to correctly identify land use and land cover based on older survey images and corresponding labels. This is where the artificial neural

Deep learning

"Deep learning" is a type of machine learning, and thus an area of the vast field of artificial intelligence. The basic principles of Deep Learning have been known for decades. For a long time, limited computing power and a lack of suitable data to feed the learning processes were major obstacles to its successful application. Once these obstacles were overcome, deep learning took off, especially in automatic image and speech recognition, achieving results that seemed unattainable until recently.

The basic principle of deep learning is to some extent inspired by the way the human brain works, which is why its algorithms are also called "artificial neural networks". However, it is ordinary computer processors that make up the hardware.

How it works

Deep-learning artificial neural networks are like vast networks of simulated neurons (artificial nerve cells), which are organised in layers and exchange signals (G2). The first layer of neurons is called the input layer. This is where the data to be examined in the artificial neural network is "stored". In the case of the aerial images for Switzerland's land use statistics, the input layer corresponds to the 120 000 colour values of the 40 000 pixels of the examined section of image (see section 4.1). Each individual colour value is thus used as the output signal of a neuron in the input layer.

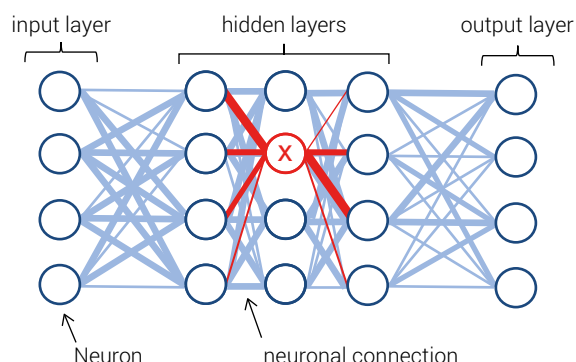
The actual data processing takes place in the hidden layers of the network. There can be a great many of these layers, making the network "deep", hence the term "deep learning". Modern deep learning networks can also have complex convolutional structures, which further enhance the performance of the network. This is also the case for the convolutional neural network employed in the land use statistics.

A given neuron "X" in the hidden layer only receives signals from neurons in the upstream layer. The aggregated signal received by neuron X will be greater or smaller depending on how strongly the upstream neurons "transmit" their signals and on the strength of the connections at neuron X. Neuron X then increases or decreases the strength of the received signal and transmits it to the neurons in the next layer. This splits the signal, so depending on the strength of the connection, a particular neuron in the next layer will receive more or less of the signal from X.

Crucially, the strength of each individual connection in the artificial neural network can vary separately, as can each neuron's strengthening or weakening of the signal. There are therefore countless small 'levers' that can be used to change the flow of signals through the different layers of the network, and thus determine which parts of the signal from the input layer arrive at a particular neuron in the output layer.

This last layer is called the output layer. In the case of Switzerland's land use statistics, the output layer neurons correspond to the individual categories of land use and cover.

Artificial neural network with three hidden layers G2



Source: FSO – Switzerland's Land Use Statistics (AREA)

© FSO 2023

It thus consists of 46 neurons for the network used to identify land use and 27 for the network used to identify land cover. The goal is to adapt the strength of the connections and the ability of the many neurons in the hidden layers to increase or decrease the strength of the signal, so that for every possible aerial image fed into the network, the neuron in the output layer that represents the correct land use or cover category always receives the strongest signal. For example, if the neural network is fed an aerial image with a sample point over a lake, then in the output layer, only the neuron that represents the use category 'standing water' should receive a signal and thus be "activated".

Learning process

The hidden layer parameters thus contain the "rules" by which the network interprets the input data. However, these rules are not fed into the neural network by a human programmer. The rules are devised by the neural network itself, hence its "artificial intelligence".

Establishing these parameters requires a lengthy training process using vast amounts of human-interpreted data and a lot of computing power. Before the training begins, all neuronal connections in the network are randomly assigned a weight, and the signal amplification or attenuation of the individual neurons is also set at random. The network is then allowed to process the training data according to these random rules. In the case of the land use statistics, the network was fed millions of sections of aerial images from previous surveys. The categories assigned by the network using the completely random parameters are of course completely useless for image interpretation. But this is irrelevant for the time being because the goal of this first pass of the data is different. Once completed, it will be possible for each aerial image to determine how much the calculated value deviates from the actual (human-assigned) value for each individual neuron in

the output layer. Based on the deviation, the network can then parametrise itself in a further, even more complex computing process, i.e. it can "learn". To do this, the network works backwards to propagate the error from the output layer to the input layer in a process called "backpropagation". The parameters that are thus established in the intermediate (hidden) layers are so extensive and complicated that humans can only partially understand them.

To speed up the learning process and increase the quality of the image interpretations, the land use statistics use a neural network that has been pre-trained by its manufacturer using other image data. Thus, when training with Switzerland's land use statistical data began, it was already able to recognise useful image characteristics such as edges or corners and was able to transfer this "knowledge" to the aerial images upon which Switzerland's land use statistics are based ("transfer learning").

network at ADELE's core comes into play. The network depends on the already much-mentioned principle of "deep learning", a special type of artificial intelligence about which you can read further details in the box on page 4. The type of network used in the case of ADELE is a convolutional neural network (CNN).

However, when the trained deep learning module is presented with a particular sample point for interpretation, the output of the computer process is not a definitive answer like "point X corresponds to use category A". Rather, the model provides an indication of probability (a prediction) for each of the 46 land use and 27 land cover categories. This result can then be used as input for further processing.

4.2 Inclusion of additional information using the random forest module

As previously mentioned, the human image interpreters can consult a range of additional information to help them interpret the aerial images correctly. The same goes for ADELE. The next stage in the process combines the output of the image interpretation process using the deep learning module with other information to further increase precision when determining the categories. At the moment, ADELE makes use of the following data:

- elevation data: altitude above sea level, slope, vegetation height, etc.
- specific data from the Topographic Landscape Model by swisstopo;
- colour values (including infrared) detected by satellite at the sample point at different times of the year;
- other geodata: official cadastral survey, building zones, utilised agricultural land, etc.

To combine the deep learning image interpretation with the additional data, a separate machine learning technique known as "random forest" is applied (see box below for explanation). Three separate models are trained:

- The first random forest model determines land use at the sample point.
- The second model determines the land cover.
- The third model only decides whether land use or land cover at the sample point have changed since the last survey (yes/no). This gives it a simpler task than the first two models, so its predictions are more accurate.

5 How ADELE is used in practice

5.1 Performance limits and countermeasures

The previous chapters clarify how the ADELE AI-tool works in theory. In practice, we can't leave the identification of the land use and land cover of all sample points entirely up to ADELE. There are two main reasons for this.

First, for many of the categories, the tool fails to achieve the precision required to ensure the consistency of past and future survey results. Despite ongoing improvements, this is unlikely to change in the medium to long term. Some categories simply look too similar, while in others – such as parks – there is a diversity of shapes and colours that the tool struggles to handle. It is therefore necessary to test which categories of land use and cover ADELE can identify with the required accuracy. The final published land use statistics can only draw on these parts of ADELE's output. In addition, the published results may only incorporate those sample points where the probability (calculated by ADELE itself) of correct identification exceeds a certain threshold. All other sample points still require human interpretation.

Second, the completely automated identification of land use and cover without a "human in the loop" is not an objective for fundamental reasons. ADELE must continue to be supplied with fresh training data in the form of human-interpreted categorisations, to:

- teach the system to use completely new or at least newly structured additional data;
- prevent certain interpretation errors from accumulating over the years in the system;
- continuously adapt ADELE's decision parameters to "physical changes" in the categories themselves. For example, "industrial and commercial areas" looked very different on aerial photographs several decades ago than they do today, and their appearance is also likely to change in future. This means that the parameters developed by ADELE on the basis of today's training data will at some point no longer be suitable for identifying this use category.

Random forest

"Random forest" is a machine learning method that is often used when objects need to be 'classified' based on certain characteristics, i.e. assigned to a specific class of objects. The method is particularly good at processing very different object features that are only loosely related. This is precisely the case with the very heterogeneous additional information (see section 4.2) used to assign the sample to the different land use and land cover categories.

How it works

"Random forest" derives its name from the "decision trees" of which it is composed. The method classifies a certain object – in this case a sample point – based on the decisions made by a large number of these individual decision trees that together make up a "forest". The ensemble model simply looks at how most of the forest's decision trees classify the object in question.

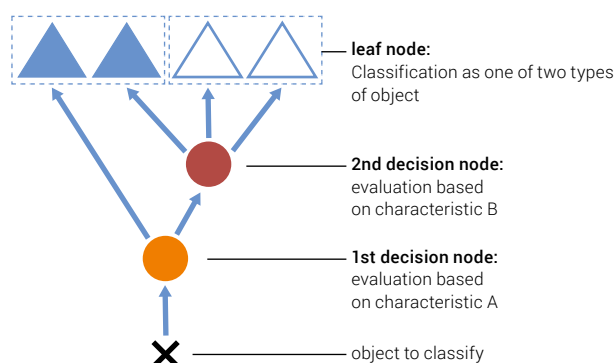
Decision trees are chains of conditional statements that can incorporate all the known characteristics of an object. For example, is a certain sample point located on a building according to the additional data from the official cadastral survey? If yes, the sample point may fall into the category "building". If no, then another characteristic of the sample point is considered, and so on. Represented as a graphic, the many ramifications of the decision chains form tree-like structures (G3), hence the name "decision trees". In the case of ADELE, a total of 280 different characteristics feed into the decision trees. This includes the 73 predictions for the 46 land use and 27 land cover categories generated by the deep learning process (see section 4.1), and 207 additional characteristics from different additional information.

Learning process

The decision trees must be trained using large amounts of data from previous surveys. Based on the available information on the sample point characteristics on the one hand and

Example of a simple decision tree

G3



Source: FSO – Switzerland's Land Use Statistics (AREA)

© FSO 2023

human classifications on the other, the branch structure of the decision trees is optimised until the trees' "decisions" are as close as possible to those of the human image interpreters. When training the many decision trees, care is taken to ensure that each branch structure is quite different from the next. A "random forest" should not resemble a monoculture of decision trees of the same type – it needs to be a mixed crop of very different tree species. This ensures a broader basis for the final classifications, which reflect all the decisions made by all the trees in the forest. To ensure "diversity", the order and thus the importance of the different characteristics included is varied in each decision tree is varied by means of random selection, which is where the "random" part comes in.

To address the second problem and thus ensure the quality of the survey in the long term, a certain number of sample points will have to be excluded from automated interpretation. The land use and land cover at these points will always have to be determined by human eye. For the land use statistics, a total of almost 1 million such points have been defined.

5.2 Procedure for semi-automated land use identification

As we have just seen, to guarantee the consistency of the statistical results, careful limitations must be placed on the use of ADELE, and its findings must be used in conjunction with the work of human image interpreters. A specific process has been

laid out for this purpose (G 4, page 6). The first stage is to check whether the sample point is one of the approximately 1 million points for which the land use and cover must be identified by a human for reasons of long-term quality assurance (see section 5.1). If this is not the case, ADELE automatically checks whether the land use and land cover at the point has changed since the last survey. To detect changes, a random forest algorithm specially trained for this task (see section 4.2) is used, based on the results of the deep learning module and the additional data. If a change can be ruled out with a high degree of probability, the categories applied in the last survey are confirmed, and the automatic determination comes to an end. If, on the other hand, there is a certain probability of change, the process to automatically detect the potentially new category is launched. However, the outcome of this process is only validated if it was ascertained

(during the testing of the accuracy of ADELE's predictions) that the probable category is one for which use of ADELE is permitted. In addition, the probability of a correct prediction by ADELE at that particular sample point must exceed a certain threshold. If these requirements are not met, the land use and land cover will still be determined by sight.

6 Findings and outlook

As of 2023, the ADELE deep learning tool can be used to determine land use and cover at about 27% of all sample points. This means that approximately 1.1 million points do not require human interpretation, which translates into a considerable time saving. It should be noted, however, that only automatic change detection is currently in use. ADELE does not yet directly assign a category when a change is detected.

In future, a number of measures will be taken to increase the share of sample points that can be automatically classified:

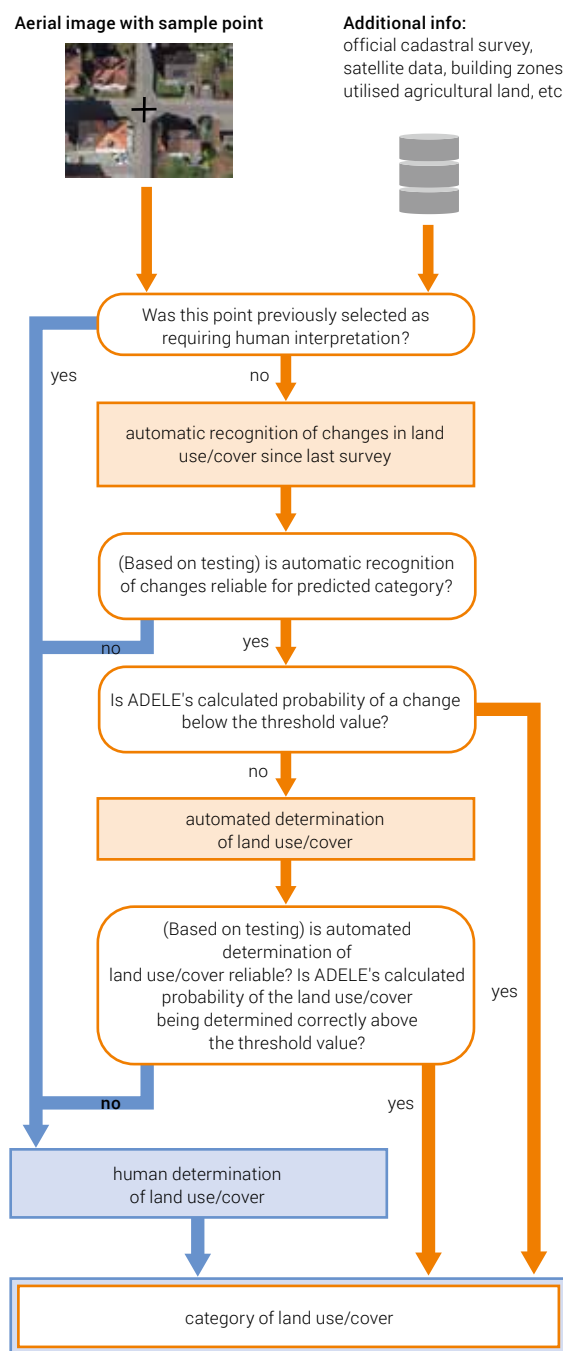
- Use of additional data: data from the Register of Buildings and Dwellings, infrared aerial images, extended time series of satellite images, land use and land cover data for neighbouring sample points, etc.
- Use of a better pre-trained neural network in the deep learning module.
- Optimisation of hyperparameters in the AI algorithms: for example, better parametrisation of the number of neurons (layers) in the deep learning module and the number of decision trees in the random forest models.

The degree of automation will also increase over time as the volume of training data continues to grow with each new survey. However, even with these improvements, it may be difficult to interpret many more than half of the points automatically one day. This limitation is also a question of resources. As the level of automation increases, further increases in automation become more and more difficult – and therefore more and more expensive. At a certain point, it becomes more efficient to allocate the resources required for further automation directly to human interpretation of the sample points. It is also important to bear in mind that the human handling of the basic data – mass viewing of aerial photos on the screen, discussions with colleagues and the field visits described above – goes beyond mere classification. Human interpretation also leads to a deeper understanding of the data and the categories used in Switzerland's land use statistics, and thus to the development of the expertise absolutely needed to interpret the statistical results correctly.

Semi-automated determination of land use and land cover in Switzerland's land use statistics: process flow

G4

The AI tool ADELE performs the steps in orange



Source: FSO – Switzerland's Land Use Statistics (AREA)

© FSO 2023

Published by:	Federal Statistical Office (FSO)
Information:	Geoinformation Section, FSO, arealstatistik@bfs.admin.ch
Editor:	Ferenc Biedermann, FSO
Contents:	Gillian Milani, FSO; Romain Douard, FSO; Ernest Hennig, FSO; Lena Willi-Tobler, FSO; Michèle Gillard, FSO; Anton Beyeler, FSO
Series:	Swiss Statistics
Topic:	02 Territory and environment
Original text:	German
Translation:	FSO language services
Layout:	Publishing and Dissemination PUB, FSO
Graphics:	Publishing and Dissemination PUB, FSO
Pictures:	Aerial photos: swisstopo; Photos: Arealstatistik
Online:	www.statistics.admin.ch
Print:	www.statistics.admin.ch Federal Statistical Office, CH-2010 Neuchâtel, order@bfs.admin.ch , tel. +41 58 463 60 60 Printed in Switzerland
Copyright:	FSO, Neuchâtel 2023 Reproduction with mention of source authorised (except for commercial purposes).
FSO number:	2264-2300

The information in this publication contributes to the measurement of the Sustainable Development Goals (SDGs).



The MONET 2030 indicator system

www.statistics.admin.ch → Look for statistics → Sustainable development → The MONET 2030 indicator system