EVOLUTION OF SE-WORKBENCH-EO FOR ARTIFICIAL INTELLIGENCE APPLICATION

Jean Latger (2), Thierry Cathala (2), Alain Le Goff (1)

(1) DGA Information Superiority, BP 7, 35998 Rennes cedex 9, France
(2) OKTAL Synthetic Environment, 11 avenue du Lac, 31 320 Vigoulet-Auzil, France

KEYWORDS: Simulation, Artificial Intelligence, Machine Learning, Infrared, Sensor, Validation, SE-WORKBENCH-EO, CHORALE

ABSTRACT:

SE-Workbench-EO, also called CHORALE in France, is a comprehensive set of tools that aim at modelling a complex and dynamic environment and at physically rendering the 3D scene for a given EO/IR sensor. Main application is focused on parametric studies. SE-Workbench-EO helps users through simulation to assess the performance of a sensor with regard to several specific environments. One important application is missile homing with EO/IR seekers.
Recently, a new application in the field of Artificial Intelligence and especially Machine Learning is emerging. The operational challenge is to improve Image Processing thanks to modern devices based on Neuronal Networks. The first step consists in preparing datasets to train the Neuronal networks and to validate the training. Today, main datasets come from real images that have to be "labelled". Synthetic images offer a key advantage to complete these datasets. The ultimate challenge is the physical realism of these synthetic images. SE-Workbench-EO seems to be a good candidate and a good starting point to take up the challenge.

The introduction shortly presents SE-Workbench-EO main features and objectives. The VV&A process will be detailed. Example of current application typically for parametric studies will be explained.

In the second part, we will discuss on the operational requirements for defence application of Artificial Intelligence and more precisely for detection/navigation application using Image Processing.

In the third part, we will focus on synthetic images datasets for Machine Learning. Advantages of synthetic imagery will be detailed e.g., automatic labelling. Associated drawbacks and limitation will be discussed.
Then the main limitation and challenge is the lack of entropy (quantity of information, lack of details etc.) of synthetic images will be explicated, both from a geometrical point of view and from a radiometric (physical material) point of view. Concerning the geometrical part, we will insist on procedural micro vegetation, micro relief and associated animations (wind …). Concerning radiometry, we will insist on amount of physical material, and resolution of the distribution of material.

Then will give information on a new approach in SE-Workbench-EO tools called "material cover" and "ground cover", that precisely bring a solution to enhance the image realism respectively for geometry and radiometry.

Finally, and before the conclusion, results will be shown.

The road map of this new orientation of SE-Workbench-EO will be presented as a conclusion.

1. SE-WORKBENCH-EO

1.1. Main objectives & features

The technical directorate of DGA has been using the SE-Workbench-EO (all versions) for 22 years to study, specify, evaluate and qualify through simulation most of defence optronic systems equipping the French armed forces. The various simulated functions comprise mainly, on sensor side, intelligence, detection/observation, homing and image processing, then, on target side, low detectability/stealth and self-protection.

Figure 1. Simulated optronic functions

SE-Workbench-EO is in fact a software that generates synthetic images in the EO/IR domain and complements a simulator that requires image generation to stimulate an EO/IR sensor and allow the evaluation of the system integrating the sensor. The relevant simulators can be divided into three families that are MIL (Man In the Loop) SWIL (SoftWare In the Loop) and HWIL (HardWare In the Loop), and lead this way to complementary requirements that the image generation must fulfil: high level of image quality, best compromise between realism and computing time, respect of equipment frame frequency.

Figure 2. EO/IR seeker HWIL trials

SE-Workbench-EO is a comprehensive set of tools that is organised into a logical and functional architecture with four main parts:

- Physical environmental modelling, which is intended to model an environmental element (terrain, vehicle, atmosphere, decoy, etc.) through the association of a geometric representation (e.g., external shape of a vehicle) and a physical (thermal, optical, rough) behaviour,
- Scenario construction, which allows to assemble environmental elements by forming a physical virtual world, define viewing conditions, create the trajectories of mobiles, manage temporal or event actions, and interactively visualize the scenario unfolding,
- Image rendering, which leads to calculating, for various sensors that observe the scene, the physical signal received by each sensor by adopting a realistic or fast calculation mode (priority is given to signal accuracy or to response time),
- Finally, the integration of sensor effects, to simulate the degradations brought by a sensor on the physical signal it receives (e.g., addition of a Gaussian noise on an image), in the same way in the realistic or fast mode.

SE-Workbench-EO was initially developed as a French government project under the name CHORALE [1]. It has become in France a reference tool shared with defence manufacturers and continuously validated with the support of ONERA. It is open for international use and is widely used inside NATO (DEU, GBR, NOR, TUR, CAN) and outside (SWE, FIN, ISR, SGP, KOR, JPN, CHN) by governmental, industrial and then academic institutes, for either defence or civilian applications.

1.2. Need for physical realism

In order to model a complex and dynamic environment, and be ready to render physically the 3D scene for a given EO/IR sensor, the first priority is to represent the target (i.e., the object of interest in the landscape) with a high level of fidelity. The target model must have fine geometrical details, an accurate mapping of observable coatings, and a detailed description of the optical properties of each coating with their spectral and directional dependencies. It can evolve over time to restore, for example, the movement of mobile parts (like tracks of a tank) or the alteration of optical properties (e.g., due to dust or moisture). In addition, for the IR domain, the model must include a precise distribution of instantaneous skin temperatures and if needed a physical representation of plume (with a possible evolution of both over time). For such a challenge, SE-Workbench-EO has dedicated modules, especially to predict realistic IR signatures of all types of civilian and military aircraft [2][3][4][5], or is coupled to other world-class software like TAIThermIR for predict the thermo-optical state of ground-based vehicles, humans and ships [6][7].

Figure 3. Tiger modelling for survivability assessment

Of course, a sensor seeks useful information about its target and must therefore discriminate between the target and the clutter. As a result, a second priority of image generation is to model the clutter with the same level of detail as the target so as not to bias the sensor operation. This priority is not really satisfied on the previous synthetic image, calculated 12 years ago, in which artefacts appear on the trees modelled each by two semi-transparent polygons (showing easily in this only way that this image is not a real one), but it has since been scrupulously taken into account [8][9][10][11][12].

Current and future optronic systems are becoming more and more efficient, able to scan ever-larger areas for longer to search for a target, using EO/IR sensors with ever-increasing angular and radiometric resolutions, and using ever-more intelligent image processing. Accordingly, this second priority applies to an ever-greater clutter and for a longer time.

In addition, potentially mobile targets cause interactions with their immediate environment. They can be mechanical, radiative, and thermal, fade more or less quickly, or persist. They belong to the clutter and must be modelled realistically, especially since they can play a major role in the detection of the target (e.g., detection of a wake and then of the ship at the beginning of the wake).

Furthermore, targets that are generally in motion seek to merge with the surrounding clutter and to be stealthy; since they are finely modelled, the clutter must also be so, as to allow to accurately characterizing the performance of discretion.

Finally, optronic scene modelling is an endless race to model ever better targets and backgrounds with ever more geometric and radiometric details with the associated animations and interactions. The synthetic image must always imitate the real image with a high level of entropy and an equivalent richness from the sensor point of view.

1.3. VV&A

Since the creation of a pole of activities for optronic scene modelling in 1998, DGA has continuously been taking care about the validity of computed images. It has been developing a general VV&A (Verification, Validation and Accreditation) process that is not limited to the modelling tools but starts with the input data for geometry and physics, and goes up to the effect on the threat behaviour [13]. Therefore, a new modelling project benefits from capitalized validation results and can focus on more demanding new test cases.

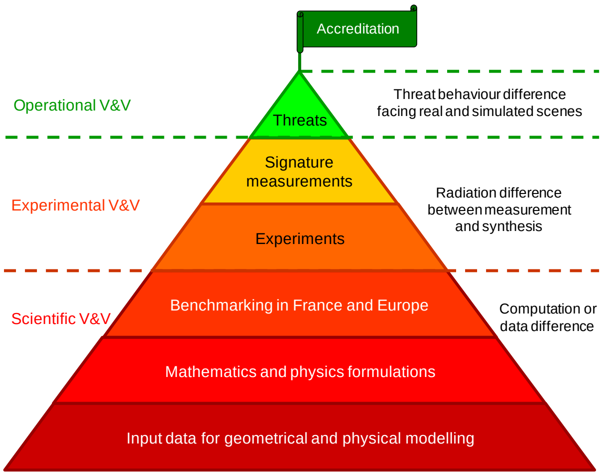


Figure 4. VV&A pyramid

The pyramid below illustrates that the tests are very simple and numerous at the base, and that their complexity increases and their number decreases when they become more and more representative, until going at the end to a few real trials in operational conditions.
In order to fully evaluate a new system, it is necessary that all the VV&A tests carried out cover at best the area of use of the system, without being out of scope, without providing a limited coverage, without suffering from sub-sampling, without showing inhomogeneity. This leads to taking into account many scenarios by varying the climate zone, weather conditions (e.g., visibility, cloud cover), time of day, target representative, target activity, nearby objects activity, type of urban or rural environment, density of buildings or plants, but also the history of what happened previously since the thermal exchanges are not instantaneous.
An interesting solution to multiply the scenarios and sufficiently sample the area of use of the simulated system is to make the scenarios parametric. As examples, the representative of the target is a parameter, the trajectory of the target is another parameter, and the thermal state of the factory in the background is yet another parameter. Thus, it is possible with a basic scenario to create with less effort a large number of scenarios by varying the different parameters.
This solution is already available and will be greatly expanded in the near future, in particular to use in a scenario some basic models that are themselves parametric, where for example the painting of a car, the presence of a trailer in the back of a truck, the camouflage pattern of a helicopter are variable parameters.

2. ARTIFICIAL INTELLIGENCE

2.1. Defence context

Defence systems are using more and more sensors to understand the environment and to better ensure the missions assigned to them. They even use sensor suites to cross-reference information and work collaboratively in a system-of-system approach to gain capacity and performance. This is true in all environments: land, air, sea, and space.

These sensors deliver, among other things, EO/IR images that are richer of information than ever before and that make the tasks of exploitation by human beings or conventional image processing more difficult. They also lead to new image enhancement treatments to facilitate the work of human beings and new automatic treatments to lighten their load. This is particularly true for detection, decamouflage and reconnaissance of land vehicles, low-level helicopters and humans.

Figure 5. Automatic detection of land vehicles

2.2. New trends

<u>Image processing evolves with AI</u>
In recent years, image-processing development has been based on conventional algorithms while looking at the possibilities brought by the AI approach. Now the strategy is different because for many future systems the AI approach makes it possible to consider increased capabilities or performance, it becomes a priority while conventional image processing becomes the backup solution.
<u>Machine Learning at a glance</u>
Artificial Intelligence, Machine Learning and Deep Learning are very general expressions that refer in essence to the capability of a machine to imitate intelligent human behaviour.

AI has been subject to several periods of advancement then stagnation, mainly constrained by lack of available data and computing power.

Actually, Machine Learning is a sub-category of AI that learns - by utilizing algorithms to parse data, learn from that data, and then apply what they have learned to make informed decisions [14].

Recently, Deep Learning has emerged as a subset of Machine Learning that rigorously mimics human brain.

Basically, the process associated to deep learning is divided in two phases:

- The machine (computer) is learning then a dedicated program is created
- The program is embedded in a system

We also say that the <u>Training</u> phase is followed by the <u>Inference</u> phase.

In this paper we focus on Deep Neuronal Networks that manipulate images. The main objective of the system is to detect and/or navigate, using images, captured by sensors embedded on a carrier, for instance a UAV or a missile, in the Defence field.

The DNN is an aggregation of unitary neurons.

A neuron possesses several entries (X) and one output (Y). Parameters are scalars that weight each entry before summation. A dedicate activation function enables to validate or invalidate the output. The validation criterion is: the weighted summation is higher than a given threshold.

This is the simpler neuron structure called *perceptron*.

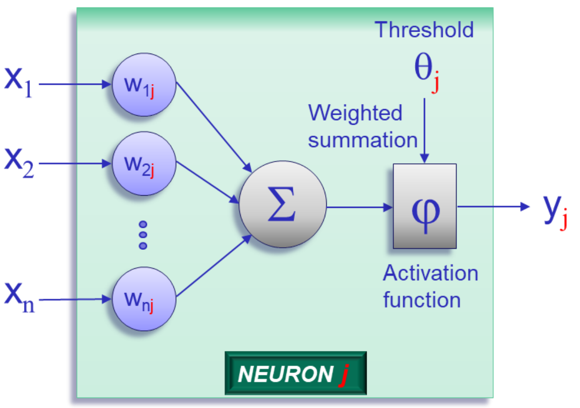


Figure 6. Neuronal (perceptron) structure

Then a generic network of "neurons" is constituted with several (Deep) layers. The DNN is born.

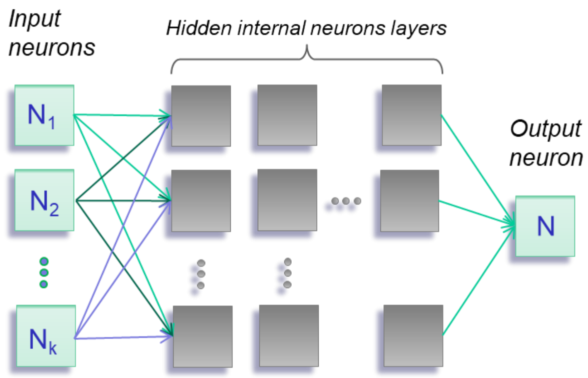


Figure 7. Full DNN network

Many images are collected. Some include the target, some not. Each image is enriched with a label. The simplest label is:" the target is present in the image" or "the target is not present in the image". Thanks to the Training phase, all combinations are tried, using many labelled (or tagged) images i.e., all the values of weights are assessed. At the end, the nicer combination that maximizes the output is kept, meaning the network is trained.

<u>Simple example illustration</u>

To give a practical example, let us take the case of the smallest perceptron using calibrated images of letters. To simplify, let us consider 3 x 3 pixels images (already filtered so that the letter is full screen, centred and not tilted). Besides, let us take the example of the detection of the "I" letter.

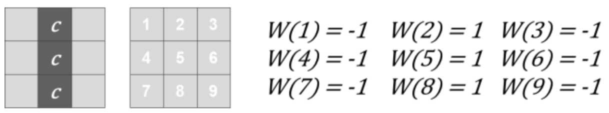


Figure 8. 9 entries and ideal weighting for "I" detection

The entry is a vector of 9 pixels either with a colour value c, either with a 0 value. The simplest network is given by the following figure:

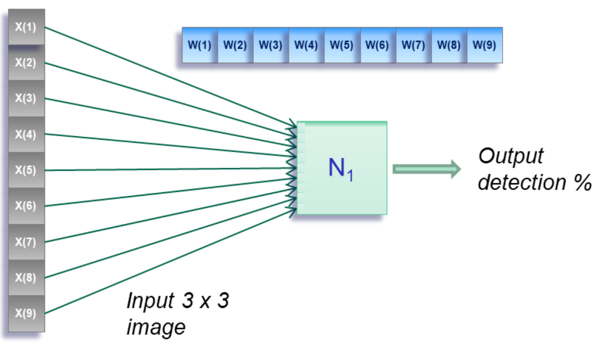
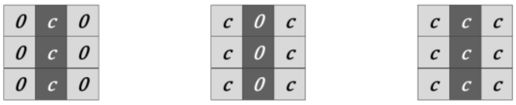


Figure 9. Simplest architecture: 1 layer of 1 neuron

The Output neuron structure is a simple weighted summation: $Y = \sum_{k=1}^{9} W(k)X(k)$

Once the network is trained, the values of the weights will be automatically identified. The higher the output is, the higher is the probability of a "I" letter.



$Output = 3c$ $Output = -6c$ $Output = -3c$

Figure 10. Example of outputs

A more complex network could be made of 2 layers of 3 neurons:

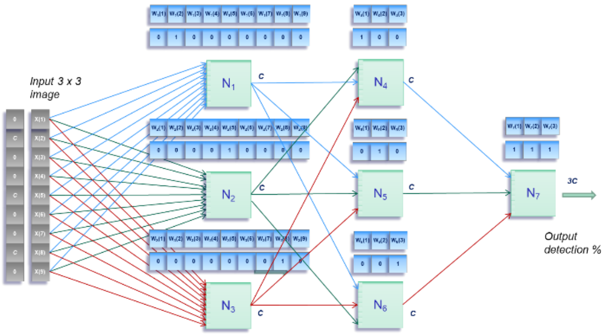


Figure 11. Architecture: 2 layers of 3 neurons

If the entry corresponds to "I" then the best score of 3*c* is obtained once the network is trained.

2.3. Neuronal Networks training

The main objective of the training phase is to design and train a model to perform a specific task with an acceptable level of accuracy.

It involves inputting sample data through the model and outputting a prediction, expressed as a percentage, that the model has accomplished its task. This cycle is repeated, adjusting the model weights by a forward and backward propagation until predictions meet the required accuracy.

<u>Classical CNN</u>
A Convolutional Neural Network is a dedicated class of NN dedicated to image processing-oriented AI models.

An intuitive look at CNN is to see it as a collection of filters [16], for instance, a filter for enhancing image contrast or detecting edges using high-pass filters.

Lower-level layers of the CNN (that are closer to the input) are made of such filters. If we remember the previous simple example to detect an "I", the convolutive layer will enhance the input image contrast, will be responsible of the input image culling (that potentially contains a "I"), to make the "I" full screen, to make the convenient translations, rotations and scales. Actually, the convolutive layer performs a *classification* that eases the job of the layers after.

<u>GAN and advantage discrimination</u>
A Generative Adversarial Network is a specific combination of 2 NN in order to improve the quality of detection, especially in the frame of imagery.

One NN is called the "Generator". The other one is called the "Discriminator". The generator's job is to generate synthetic images as realistic as possible and similar to the training data. The job of the Discriminator is to detect whether the image is real or fake. The Discriminator feeds back its output into the Generator to improve the cycle in case of fakes until the NN does not make any mistake. The more the Discriminator is sensitive, the more the images generated by the Generator are realistic.

The well-known *pix2pix* application uses GNAs for image conversion:

- Low resolution to higher resolution
- Black & white to colour
- Sketched into realistic.

GANs have been widely applied to generating realistic medical images. Another application is Arts and paintings for imitating famous artists. More recently, GANs have been used for "Deep fakes", for instance to mimic politicians.
GAN is the technics behind what is often called "Style Transfer" (for example for imitating Van Gogh painting style).

An interesting application concerns the mimicking of the Transfer Function of a sensor, typically an IR sensor, especially if the characteristics of the sensor are not precisely known.
Nevertheless, if the sensor is known, we consider much more efficient to use synthetic images that are computed with a scientific model of sensor (SE-Workbench).

Anyway, GAN constitutes a very promising means to make synthetic images much more realistic in order to populate the training datasets.

<u>Transfer learning</u>
Transfer learning consists in reusing an existing AI system that has started to learn from a given dataset for a given class of object detection in order to apply it to extended classes of objects. In other words, it aims at removing only the last few layers of a given NN and reutilizing the generic layers. More simply, it is making new with old.

Transfer learning is very important for industrial application, when the Training and Inference phases are very expensive. It enables to expand the application domain of an existing and pre-validated model.
But it is very important to understand that TL is not magic. Many PhDs in France and over the world investigate "impossible" TL such as transforming a NN trained for visible domain images into RADAR imagery or THERMAL imagery.

In that case, a much more efficient way, is to use synthetic images, based on a physical approach, to enrich the training datasets.

Frugal Learning
In Machine Learning domain, we distinguish Active Learning and Incremental Learning. AL is used in cases where data are available but where labelling is quite expensive. By opposition, IL consists in continuously training the algorithm. IL is an "On Line" training when AL is an Off Line". Frugal Learning mainly concerns IL. FL is a means to overcome the lack of data and/or overcome the cost of data. In simple words, FL is a specific ML declination with minimal resources.

Typically, in the military domain, a complete database of targets is rarely available, since confidential. In that case, FL becomes mandatory.
FL aims to build the most accurate models as possible using the least amount of labelled image. The basic FL method consists in scoring then filtering the training datasets.

Synthetic images represent a nice opportunity for frugality. We can start with a huge amount of synthetic image and progressively simplify the dataset until the predictive accuracy of the NN remains good.

Thanks to synthetic images, it is possible to:
- Compute all the possible combinations of images (target position, orientation, scale, vicinity, movement…),
- Compute, for a given image, a set of Levels Of Detail of this image with several image qualities.

Synthetic imagery is a good candidate for reducing image quantity, focusing on quality and selecting good images.

2.4. Neuronal Networks assessment

Make the NN explainable
Explainable AI is not a simple AI wherein the model only provides predictions. Explainable AI also accounts for the factors that triggered a given prediction and reveals its limitations. To make these models more explainable and interpretable, the *heatmaps* come to the rescue. A heatmap emphasises and shows a part of the image that leads to the prediction with higher intensity. For instance, it is interesting to know that datasets for recognizing helicopters are more sensitive to the windscreen than to spinning blades.

Make the NN reliable
Explainable AI is a means to fight bias and focus on good images (frugality). A classical tool to assess the robustness of the NN is the test of *occlusion sensitivity*. The idea is to mask a part of the image (e.g., using a small rectangle placed randomly over the image) and measure the prediction rank.
Besides, many research centres and SMEs have worked a lot on *formal validation* of AI, especially in France. Formal validation tools have been developed.

These tools generate a dedicated heatmap that maps the formal validation ratio as a safeness percentage. A complementary method consists in introducing noise and tracing the curve robustness as a function of the noise power for several types of labelling.

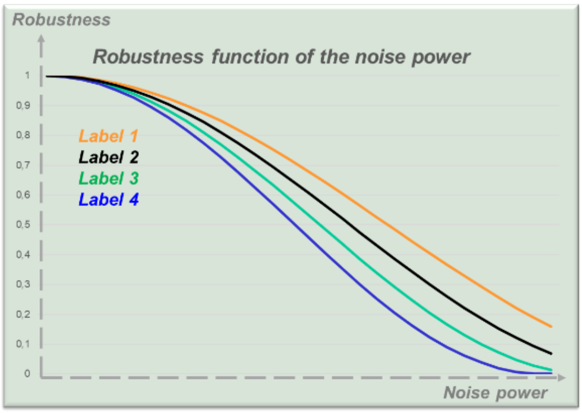


Figure 12. Robustness assessment

3. SYNTHETIC DATASETS FOR AI

In various domains, deep learning algorithms are exploited to process images data. Several applications are AI demanding:
- Earth observation from space
- Security surveillance (civilian safety)
- Guidance systems
- Vision based navigation systems (automotive, aviation, UAV, maritime and railway industries)
- Detection, Recognition and Identification (DRI) systems, in the defence industry

In these applications, and especially in the defence domain, the access to relevant training database is not ensured. In the defence field, the sensitivity of information leads to a noticeable scarcity of real learning data. The recent increase of sensor fusion systems merging visible, IR and even radar images to enhance the detection capability makes things harder when comes the time to gather learning data covering all the sensor images of the same area at the same time.

The huge amount of image data available in world databases like COCO or ImageNet are most of the time in the visible domain only and cannot be considered as a solution for the critical defence applications.

3.1. Why synthetic images

Dedicated approaches are currently investigated to tackle the lack of relevant real data. For instance, neuronal style transfer could help to create consistent IR data as seen previously. But this becomes harder when considering Radar images (like Synthetic Aperture Radar) where there is no obvious correlation with an electro-optical image taken by a camera.

Frugal learning techniques are also investigated but still rely on the availability of some few real labelled data, which could be really challenging when addressing confidential areas or targets in the defence domain.
Therefore, creating learning data, as we would expect the real ones to be, appears very attractive and anyway the only remaining solution.

3.2. Advantages of the synthetic approach

Many advantages come from a synthetic approach for AI:

- **Automatic Labelling** (tagging)
  The following figure illustrates 2 basic labelling that are today available in SE-Workbench-EO, one gives every object type and apparent surface, the other one gives the distance of the pixel to the sensor:

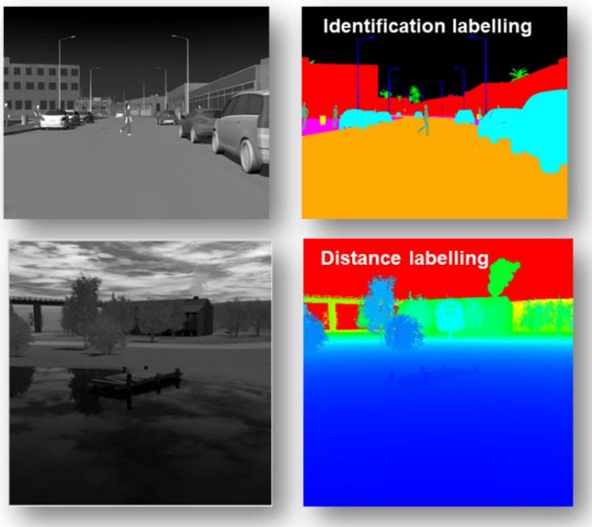


Figure 13. Metadata associated to SE-WB images

  Many other automatic labels are on the road in SE-Workbench such as:
  - Behaviour (altercation, weapon…)
  - Pixel radiometry
  - Masking
  - 2D/3D bounding boxes
  - Intermediate sensor output
- **Knowledge of the ground truth**: labelling is straightforward, no need for a skilled photo interpreter
- **Multi sensors capability**: same time, same position, same line of sight for any waveband
- **Variability** and situation diversity: ability to consider any environment (background, weather, lightning conditions, clouds, targets, flares …)
- **Repeatability**: ability to carry out parametric studies
- **Big data capability**: No limit in the number of images.

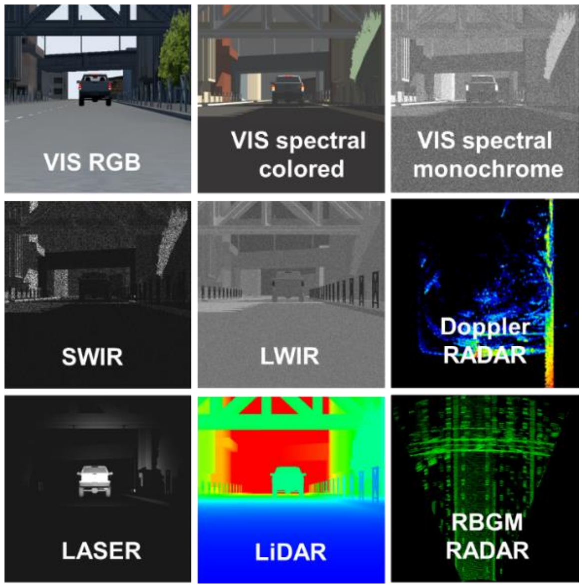


Figure 14. 9 consistent sensor simulation using SE-WB

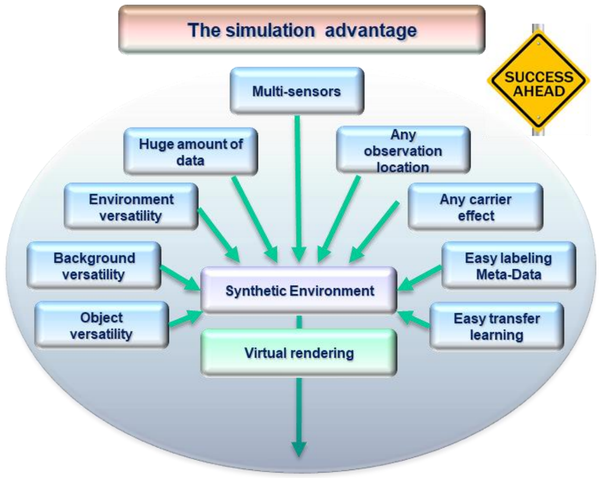


Figure 15. Main advantages of simulation for AI

3.3. Drawbacks of the real datasets

The three main problems due to real images training datasets are the following:

- The **lack of real usable images** especially for sensitive application (e.g., defence) that is more and more demanding for sensors
- The **risk of coherency** of images in case of sensor fusion. It is almost impossible to find real images of same scene/time/orientation from both EO/IR and RF sensors
- The **cost of labelling** and the risk of errors and inconsistency.

Other drawbacks can be noted such as:

- **Overfitting**: due to limited size of training data sets to train an algorithm
- **Unbalanced training data sets**: lack of diversity
- **Observation locations**: difficult to find real images of the same scene from different viewing angle
- **Tagging quality**: tagging meta data in the images is limited

- **Repeatability**: impossible to change a single parameter at a time in a real image
- **Transfer learning**: difficult to turn a dataset dedicated to one sensor into a dataset suitable for another sensor. This is possible with SE-Workbench-EO. The perfect image before sensor is archived and can be reused for several sensor models.

3.4. The good balance?

Actually, the main drawback of synthetic images is the lack of realism. The objective is not to train an AI system that is embedded in a simulator but to use it in the real world. Of course, the real data does not suffer from a lack of realism.

As a matter of fact, an intermediate approach is to use the available real images then exploit synthetic images as an add-on to complement those real images.

A more sophisticated approach, implemented by DGA, is the recombination of images to densify existing datasets (real and synthetic images) and make studies more robust, the field truth being naturally produced during the incrustation.
Simple recombination is the data augmentation with initially consistent data: this allows to copy/paste, move and multiply elements (e.g., targets) in an image by applying control metrics. That can be done with either real or synthetic images.

Figure 16. Augmentation with a tank then a tree coming from the same series of images

A more complex recombination is the hybridization of initially inconsistent data with an ability to improve consistency (scale, sensor, distance, contrast, colourimetry ...). This also makes it possible to embed synthesis objects in real backgrounds, which make it possible to limit study bias.

Figure 17. Hybridization of SE-WB target and real background with various contrasts

3.5. Requirements for synthetic images

During the last three years, OKTAL-SE has computed many datasets for many companies, research centres and state agencies, especially for defence application (e.g., Man Machine Teaming project [19]). The lessons learned from these experiments is clear:

- The synthetic data have to be **realistic**, deep-learning algorithm wise
- The training process and the CNN configuration have to be **adapted to synthetic data**

3.6. Avoid traps with synthetic images

Good looking of synthetic images only, this is not sufficient for warranting the consistency with real world.

SE-Workbench-EO has recently made significant progress for competing with reality.

Typically for targets representation:

Figure 18. Visible and IR rendering of SE-WB targets

But also when integrated in the environment:

Figure 19. Influence of lightning and weather

And it is now possible to generate geo-specific images with SE-Workbench-EO which are very difficult to distinguish from real ones:

Figure 20. From reality to simulation

More and more visible cameras that are very useful for detection and navigation functions include SWIR (Short Wave Infrared) capabilities. It is obvious considering mobile phones cameras that are able to take good pictures by night. Many VIS-SWIR cameras are now available on the market.

IR simulation implies to work in the spectral domain. SE-RAY-IR tool is the ray-tracing kernel of SE-Workbench-EO. SE-RAY-IR is a spectral ray tracing. Every feature is spectral: materials, atmosphere, and detectors. SE-RAY-IR manages the combination of these spectral features with respect to the physics laws. It computes hyper spectral images (roughly 10 to 100 wavelengths in the sensor spectral band) as input for the sensor then SE-IR-SENSOR module turns these radiometric images into an after-sensor signal. So, IR synthetic rendering implies to work in the spectral domain. Even for the visible band that is often extended by SWIR capabilities, it is also mandatory to work in the spectral domain.

Video games images are RGB images. A pixel is green because it has been "painted" in green colour. Deducing its "colour" in the IR domain is quite impossible. Considering SE-RAY-IR, a pixel is green because its reflection factor (in fact the Bidirectional Reflection Distribution Function) - that is typically defined from 0.2 micron to 16 microns - is particularly high around 0.54 micron i.e., the green wavelength.

In SE-Workbench, the sun and the sky dome emission are spectrally defined. The ambient light is modelled as a white light (slightly yellow) nearly constant over the visible spectrum. So, the green colour is due to the coupling of the solar spectrum and the "green" reflection factor.

Consequently, RGB video rendering cannot be seriously used for simulating large band sensors such as VIS-SWIR cameras. Spectral computation is not an option, it is mandatory.

Another common mistake is to assess the quality of an image using human eyes. Actually, it is not applicable for any application in the technological world, it is not a human being who sees but a hardware device. An Advanced Driver Assistance System is a good example in the automotive domain.

In the IR domain, part of the quality of the image is due to its spectral content but also to the number of bits (resolution depth). A 16 bits image that provides 64k grey levels is much more efficient for detection than a black & white image limited to 8 bits (up to 10 bits in new GPUs) and coming from video game. This is the reason why SE-RAY-IR computes radiances (W/m^2/sr/m) and not colours in floating point precision.

Synthetic images for AI have to respect at least:

- Computation in the spectral domain
- Computation in double precision
- Add-on of the Sensor Transfer Function

More conceptually, we have to fight against:

- A common bias about realism perception:
  - Is realism judging "human eyes" or "algorithm perception"?
- Physics and technology modelling is not an option:
  - It is vain to mimic reality and focus on cosmetic images, even in the visible domain, even if you have two eyes

4. SE-Workbench and AI

4.1. AI for SE-Workbench

<u>Source data cleaning & improvement</u>

One important source of information that precedes the synthetic environment modelling is ortho-images. SE-Workbench includes the SE-AGETIM suit of 3D terrain modelling tools that automatically shape the terrain, profile infrastructures (roads, rivers…) and extrude superstructures (trees, buildings…). The real add-value of SE-AGETIM is the management of physical materials and the preparation of the scene for visible, IR and radio frequency applications.

Ortho-images carry lots of useful information for terrain generation. For the terrain, most of the time, orthoimages are mapped onto the terrain

tessellation. Nevertheless, there is a lot of artefacts. For instance, shadows. Shadows are not intrinsic. It is the job of the renderer to compute shadows (with a physical approach and depending on the ephemerid). So, we have to rub and erase them from the ortho-image. More generally all 3D objects, projected on the ortho-image have to be suppressed. For instance, in the case of a building, it is important to replace this 2D artefact information by a 3D object. In RGB approximation, it is not so critical. Of course, in IR (due to thermal computation) and in Radar (due to dihedron and trihedron effect), it is mandatory.
A special case is the ephemeral object. For instance, a car onto a highway will be simply erased. A car onto a parking will be replaced by a 3D instance.
AI is a perfect vehicle to perform this clever cleaning.

Figure 21. Example of ortho-image cleaning

Another important support from AI concerns the ortho-images improvement. The first need is to enhance the resolution of existing ortho-images adding “noise” and/or re-synthetize the image (GAN approach). The second one is the colour harmonization. It is a complex problem that is very hard to perform manually. Actually ortho-images are very depending on the acquisition sensor, on the cloud cover and on the date and country.

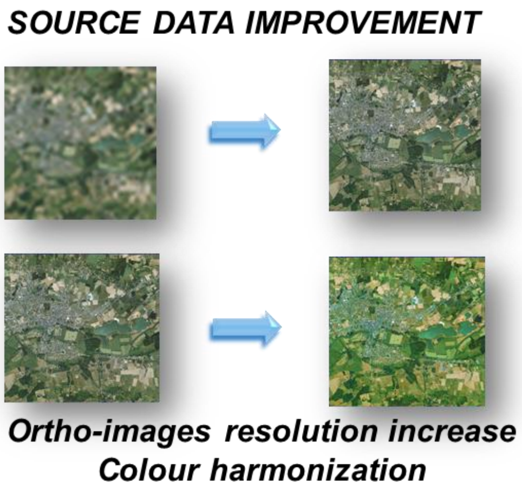


Figure 22. Example of ortho-image improvement

AI is also a perfect vehicle to perform this clever improvement.

Automatic physical classification

Classification consists in selecting the good physical material in a 3D synthetic environment. Today, SE-PHYSICAL-EDITOR is a manual tool, slightly assisted by some Image Processing algorithms, that is used by graphists in order to assign the proper reference to the physical material database provided with SE-Workbench.
To do this human classification, all types of information are welcome. The main one is an image. The goal is to recognize what it is made of.
This operation is obviously made for AI and will be all the more efficient as the amount of data increases.

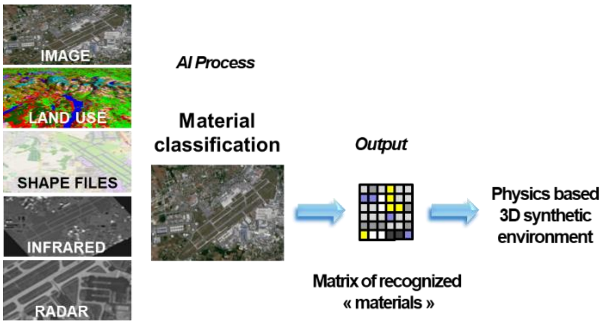


Figure 23. Using AI to automate and enhance physical materials classification

4.2. SE-Workbench for AI

The new trend for SE-Workbench is to provide synthetic datasets for NN training.

Challenges to address

The key challenge is the image entropy or variance. Entropy definition is based on statistical mathematics when variance concept is based on classical Image Processing.
The main risk with synthetic images is the lack of entropy. In practise, we can figure out the entropy concept using simple examples.
The first component is the geometric detail. Recent progress of computer graphics eludes this problem. Last decade, millions of polygons constituted a real challenge. Today, billions of polygons are reality. The second component is the lighting model. Thanks to ray-tracing techniques, shading, global illumination, ray marching in voxels and participant media provide nice solution. Then, Particles Systems for smokes, animatics for vehicles and skeleton animations of characters participates efficiently to realism. But the key feature is the entropy due to materials that are referenced through textures. A great amount of intrinsic physical materials is mandatory. Then complex combination of these materials within a texture conveys high variance level to the images.

Material Cover & Ground Cover

One recent improvement of SE-Workbench is **Material Cover**. To increase the texture resolution, the idea is to invent details inside a given texel (the texture element). Let us consider a 1024 x 1024 texture covering 1 km x 1 km. The texel size is then 1 m x 1m. In SE-Workbench, the texture does not contain a value but a reference to a physical material that contains lots of values. Each category of material is referenced in the 1024 x 1024 master texture. For instance, some texels reference earth, other gravels, other grass etc. The Material Cover idea is to replace the master texel by a second order new texture. For instance, the 1m x 1m texel of gravel is replaced by a 256 x 256 texture that represent an aggregation of asphalt and several varieties of stone. Then the resolution of the second order texel is 4 mm x 4 mm.

Figure 24. Material cover used to enhance detail of a 5mx5m ortho-image of an airport, rendered by SE-WB

Material Cover is a 2D super sampling. SE-Workbench allows a more complex 3D procedural enhancement, typically for grass or small vegetation, called **Ground Cover**.

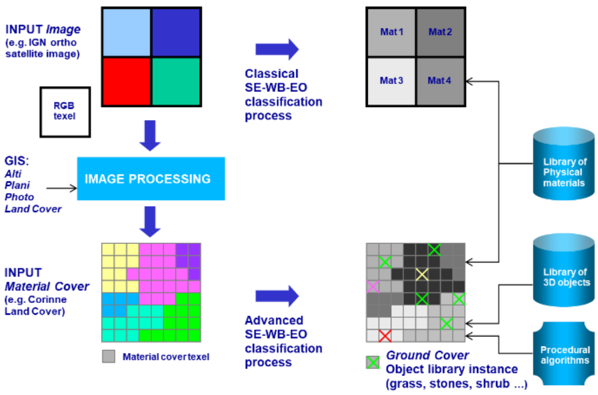


Figure 25. 3 ways to use texture in SE-WB: Classical classification, Material Cover & Ground Cover

Ground Cover is very efficient for increasing the entropy of synthetic images:

Figure 26. Ground Cover to figurate an assembly of grass and stones in SE-WB

Wang Tiling 2D & 3D

Another important improvement in SE-Workbench is the **Wang Tiling** algorithm. To remain simple, let us say that WT can transform a texture into a quasi-infinite set of similar textures, ensuring continuity between these textures, but without any repetition. WT protects against the "ravioli" self-repetition effect that does not exist in the real world of course.

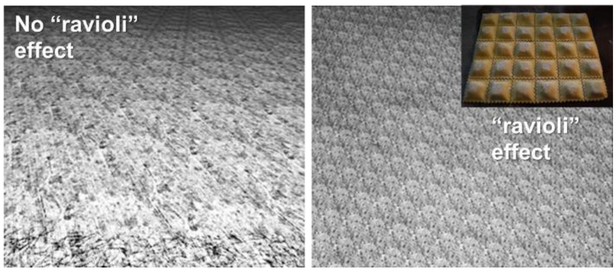


Figure 27. WT in the Left image: a means to fight against repetition and improve image variance

OKTAL-SE has generalized the WT algorithm to 3D, pushing the instantiation of similar 3D features but never self-repeated. It is a very efficient approach. First the modelling cost is very light since limited to a few templates. Then, no limit in details refinement since automatic. Finally, no room for error, neither for geometry nor for physical classification since the manual operations are reduced to a very limited set of templates.

Figure 28. Thousands of buildings for a dozen templates using WT 3D automatic generation

Figure 29. Millimetric details available for short range observation using WT 3D

5. DISCUSSION

Using physics-based sensor simulation software to generate synthetic datasets is a promising concept. However, some challenges have to be faced to turn this concept from a promising idea to an operational solution:

- A dedicated process, adapted to synthetic data, has to be developed for the training and the assessing phases of a CNN
- The quality of the generated data has to be improved (more Physics, more technology in the sensor modelling, more details in the 3D scenes…)
- More sensor modelling capability, in order to address fusion and take advantage of transfer learning and data frugality
- Tighten partnerships with partners for progressing (GAN, domain randomization, explanation, robustness, formal validation)

6. CONCLUSION

Neuronal Networks synthetic training is an exciting approach for future. It is not today efficient enough due to the lack of physical realism of synthetic rendering tools. In that domain, SE-Workbench constitutes a key means of generating trustable synthetic images since SE-Workbench mimics Physics when many “cosmetic” tools do not. OKTAL-SE will pursue its long road in order to make SE-Workbench closer and closer to reality.

7. REFERENCES

1. Le Goff, A. & Latger, J. (1999). Realistic multi spectral simulation including IR simulation, *SPIE Proceedings*, Vol. 3694, April 1999

2. Le Goff, A. (2010). Advanced modeling and rendering of stealth aircraft optical signature, *OPTRO-2010-1818382*

3. Le Goff, A. (2010). Predictive modeling of aircraft IR signature, *ITBM&S proceedings*, 2010

4. Le Goff, A. (2016). Validation effort with PRESAGE for aircraft infrared signature prediction, *OPTRO2016-061*

5. Le Goff, A. & Kersaudy, P. (2018). Strategy to model the thermo-dynamical cycle of a gas turbine from open source data, *OPTRO2018_072*

6. Le Goff, A. (2014). Thermal modeling coupling RadThermIR targets and SE-Workbench environment, *OPTRO-2014-2966329*

7. Le Goff, A., Hurtaud, Y., Floch, E., Corbihan, P. & Jau, C. (2018). CUBI experiment & simulation for infrared scene modeling validation, *OPTRO2018_037*

8. Le Goff, A. (2012). High realistic IR terrestrial scene modeling for intelligence function assessment, *OPTRO-2012-106*

9. Le Goff, A., Cathala, T. & Latger, J. (2015). New impressive capabilities of SE-Workbench for EO/IR real time rendering of animated scenarios including flares, *SPIE Security+Defence 2015-9653-6*

10. Latger, J., Pajot, A. & Cathala, T. (2018). Recent improvement of the “Fast” version of SE-Workbench-EO, *OPTRO2018_00053*

11. Cathala, T. & Barbé, S. (2018). Validation of SE-Workbench-EO in the visible spectral band, *OPTRO2018_00002*

12. Le Goff, A. & Fèvre, JM (2018). Defense applications of maritime scene simulation with SE-Workbench-EO, *ITBMS 2018*

13. Le Goff, A. (2008). IR scene modeling VV&A, *ITBM&S 2008*

14. Zendesk (2017). A simple way to understand machine learning vs deep learning. *https://www.zendesk.comSmith*

15. Le Cun Yann (2020). When the machine learns. *Odile Jacob*

16. Koul, Ganju, Kasam (2020). Practical Deep Learning for Cloud, Mobile & Edge. *O’REILLY*

17. Fiammante Marc (2019). History and uses of the newly published artificial images for neural networks patent. *LinkedIn*

18. Tremblay Jonathan (2018). Training Deep Networks with Synthetic Data: Bridging the Reality Gap by Domain Randomization

19. Man Machine Teaming project, in partnership with Dassault Aviation and Thales companies. *https://man-machine-teaming.com/oktal-se-magellium-et-le-projet-constitution-de-bases-de-donnees-de-scenes-synthetiques-eo-ir-rf*