1. Introduction

LTU delivers visual intelligence solutions for image management and recognition tailored for your use case.

Every case is considered individually. Depending on your need, LTU builds the appropriate scenario by calling multiple algorithms. These main features are packaged in LTU Engine which provides the components necessary for creating and managing visual search applications, including JSON API and a comprehensive Administrative Interface. Each technology provided and used by LTU is explained in this document.

The solution is available via licensed software or via the hosted platform. You can choose to deploy it on your own servers for security or privacy reasons, or use LTU Cloud solutions if you don’t want to deal with server purchasing and maintenance.

2. Image management

LTU Engine is a complete solution that will help you to organize and manage your images. All the following solutions could be used to structure and manage your images:

- **Indexation:** to create your private database of searchable images
- **Deep learning:** to classify your images and to get associated keywords
- **OCR:** to get text informations from an image
- **Colour algorithm:** to get colours of an image

Once your images are structured, you will be able to search inside by using keywords or/and image search algorithms.
2.1. Image indexation

LTU provides image recognition solutions that allow to find references from a query in a private image database. The search solutions are described in the second part of this whitepaper.

The first step towards making an image searchable is to create a descriptor of the image content. LTU Engine computes a visual signature for every image that describes its visual content in terms of color, shape, texture and many higher order visual features.

These descriptors are also called image DNAs. The DNAs are unique for each image and specific to a search algorithm.

![Image indexation diagram]

LTU Engine stores the DNAs in a database that constitutes your reference image database, in which you would search by using LTU query retrieval solution. LTU Engine is fully optimized and let you index a collection of millions of images in a private database, store on a standard server/computer and run all kinds of queries on it in the twinkling of an eye.

2.2. Artificial Intelligence

LTU uses deep learning models to classify large batches of images, detect objects in an image and generate keywords. Our computer vision processes aim at providing our clients with bespoke, fast and accurate image recognition.

LTU detects retail products, decorated objects, pictures, book covers, textbooks, art paintings, logos, and more focused on Image processing.

2.2.1. Deep Learning and Transfer Learning

Deep Learning is a type of artificial intelligence derived from machine learning where the machine is able to learn on its own. Deep Learning is based on a network of artificial neurons inspired by the human brain. This network is composed of tens or even hundreds of "layers" of neurons, each one receiving and interpreting the information of the previous layer.

At each step, the "bad" answers are eliminated and returned to the upstream levels to adjust the mathematical model. As you go along, the program reorganizes the information into more complex blocks. The model is pre-trained on a training dataset.

When this model is subsequently applied to other cases, it is normally able to recognize an object without anyone ever teaching the concept of this object.

![Deep Learning concept]

Dee Learning can be optimized thanks to Transfer Learning; the knowledge acquired from the training dataset, which is called the "source" dataset, is "transferred" in order to properly handle the new dataset, named "target". For example, knowledge gained while learning to recognize cars can be used to some extent to recognize trucks. Thus, Transfer Learning allows us to build accurate models in great time efficiency.

Based on the Transfer Learning principles, LTU implements image classification tailored for your specific needs as described below.
2.2.1. Detection and Identification

Image Classification is the task of taking an image as an input and outputting a class label from a set of classes to which the image belongs.

The process normally involves recognition of the dominant content in a scene. The dominant content gets the strongest confidence score irrespective of the transformation of that content such as scaling, location or rotation.

LTU Engine Interface allows you to drag and drop your dataset and annotate each image using your specific terminology.

Choose state-of-the-art Deep Learning algorithm:
- Single Shot Detection (SSD)
- You Only Look Once (YOL0)
- Faster R-CNN

Train & Test

High Accuracy & Real-time Object detection

2.2.2. Keywords

As explained, Deep Learning helps LTU to classify your images in pre-trained classes. LTU also allows to associate these classes in keywords. Thus, Keywords will help to organize your dataset, to search images with keywords inside your dataset and to orient the image search process.

From output of object detection & classification

Label image using keywords from detected & classified objects

Cluster similar images in database based on keywords proximity

2.3. Optical Character Recognition

Optical character recognition refers to both technology and process of reading and converting typed, printed or handwritten characters into machine-encoded text or something that the computer can manipulate.

LTU uses OCR solutions to extract text from an image. The retrieved information could help to:
- Classify images dataset
- Determine keywords

Logo
Vogue

OCR
Rihanna / Fashion / Style / Paris

Colour Palette

Keywords
Women Red Paris
Rihanna Fashion Black
2.4. Colour palette

2.4.1. Image colours

LTU Engine can return you the list of colours that are present in an image.

2.4.2. Trend of an image collection

LTU Engine Colour can analyse an image collection and return the most frequent colours. The set of the most frequent colours is what we call a colour palette.

The colour palettes can be used to:

- suggest relevant queries to the user
- provide a quick overview on an image collection

An interesting feature of the palette is that they can be computed on any subset of a collection.

Subsets can be categories. For instance, LTU Engine can compute a palette for the "Women Shirt" category. This will be different from the whole images collection palette. Some colors that are not present in this category will be removed and LTU Engine will introduce color nuances for the most present colors.

These subsets can also be result sets. If they are used to propose queries to the user, this feature can be a powerful tool for query refinement.

3. Image recognition

LTU provides image recognition technologies via its product LTU Engine. The solution is available via licensed software or via the hosted platform: LTU Engine OnPremise/OnDemand.

LTU Engine includes two distinct images processing technologies:

- The Visual Search that is divided into two recognition solutions – image matching and visual similarity search.
- The Image Processing that offers a Fine Image Comparison solution.

3.1. Visual Search Solutions

3.1.1. Overview

The visual search solutions allow to find, from a query image, identical or similar visuals in images databases. The search is based on object recognition, shape or color and depends on the content of an image rather than textual information.
Our clients use the visual search for:

- **Art Identification**: to know if an art work is stolen
- **Brand Intelligence**: to survey if a merchandise is not counterfeited
- **Media Intelligence**: to analyse what the Internet relates about a brand or a product
- **Place Detection**: to find a place from a picture
- **Page Identification**: to get information of a product from an ad in the street
- **And more…**

The visual search is composed by two key steps:

- **Indexation**: *As with the reference images*, the first step is to create a descriptor of the image content. LTU Engine computes a visual signature for every query that describes its visual content in terms of color, shape, texture and many higher order visual features. These descriptors are also called image DNAs.

- **Retrieval**: A special comparison technology by which an image signature can be compared at a extremely high speed to other image signatures from a database up to millions of images.

Each search returns a reference list, their distance (or score), optional keywords as well as additional algorithm details.
### 3.1.2. Visual distance

The distance is an indicator for the relevance of the retrieved images: the closer the value is to 0.0, the closer the retrieved image shares the same visual content as the query image.

<table>
<thead>
<tr>
<th>Distance</th>
<th>Matching/Color</th>
<th>Similarity</th>
<th>Irrelevant</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Clone</td>
<td>0.1</td>
<td>Poor</td>
</tr>
<tr>
<td>0.1</td>
<td>0.9</td>
<td>1</td>
<td>Good</td>
</tr>
<tr>
<td>1.0</td>
<td>1.5</td>
<td>1.8</td>
<td>Poor</td>
</tr>
<tr>
<td>1.8</td>
<td>2</td>
<td>2.2</td>
<td>Irrelevant</td>
</tr>
</tbody>
</table>

The visual distance is normalized such that a value:
- equal to 0 is a clone
- below 1.0 indicates a match
- between 1.0 and 1.8 reveals a similarity

---

### 3.1.3. Image Matching Technology

#### 3.1.3.1. Overview

The image matching technology is used to find, in database(s), images that:

- Look exactly the same (e.g., for deduplication)
- Have been edited in any way (e.g., for tracking on copyright images)
- Are photos taken of the same visual content (e.g., for print to mobile applications)

---

### 3.1.3.2. Image Transformations

LTU Engine's image matching technology is robust against several types of image transformations, detecting not only the exact same image, but also modified versions of the original image and object matches (photographs of same object).
This part illustrates the types of image transformations that LTU Engine can handle in order to identify a match. Image transformations can be broadly divided into several groups:

- **Geometric transformations**: Scale changes, rotations, translations, flips and projective transformations.
- **Photometric transformations**: Colour space conversions, gray level transformations, changes in hue, brightness and contrast.
- **Filtering effects**: Noise and smoothing.
- **Structural transformations**: Composite images, i.e. images that have been overlaid on top of each other, overlay of text, adding of borders and cropping.
- **Recompression**: Due to different compression algorithms (such as jpeg) and different image encodings, information can get lost and artifacts may appear.
- **Images derived from mobile**

Images will often be modified with a combination of the above transformations. The LTU image matching technology is robust even in these instances. The matching technology easily matched the above combination which includes Gray scale, blur, re-encoding, projective transformation and overlay composite transformations.

### 3.1.3.2.1. Geometric Transformations

LTU Engine is capable of identifying image matches despite geometric distortions.

- **Resizing of the original image**

![Original](image1)
![Resized](image2)
![Resized, change in aspect ratio](image3)

- **Arbitrary Rotations**

![Original](image4)
![Rotation -37°](image5)
![Rotation 12°](image6)
![Rotation -30°](image7)

- **Projective distortions**

LTU Engine matching technology is capable of handling some degrees of perspective distortions.

![Original](image8)
![Negative Rotation](image9)
![Positive Rotation](image10)
3.1.3.2. Photometric Transformations

LTU Engine's Image matching technology can detect matching images regardless of these photometric transformations:

- **Grayscale**: Colours of the image converted to shades of gray.
- **Brightness**: Luminance settings correspond to the degree of luminance within each image pixel. For a distant observer, the word ‘luminance’ is substituted by the word ‘brightness’, which corresponds to the sparkling parts of an object or image.
- **Contrast**: The difference between the darkest and the brightest parts of an image.
- **Color change (Hue)**: Changes in coloration, hue is a complex color obtained by a mix of basic colors - Red, Blue, Green.

3.1.3.2.3. Image Filtering and Noise

Filtering effects are mainly linked with image printing, but also with modifying image metadata. Filtering transformations affect the image ‘clarity’. Depending on the filter used, they can either sharpen or blur the image. LTU Engine's image matching technology processes these images without difficulty.

3.1.3.2.4. Structural Transformations

Changes related to structural transformations affect the structure of the image. These transformations do not limit the matching of images.

- **Framed, flipped, text added, cropped**: Also, LTU Engine matching technology is capable of handling:
  - **Addition of a border or frame**: A border of uniform color is added on one, several, or all sides of the image.
  - **Flip**: Using a particular configuration of LTU Engine’s image matching signature optimized for image tracking applications, the technology is capable of matching flipped images.
  - **Addition of text to the image/superimposition**: The addition of text to the image with or without a background. With LTU Engine's technology images are matched regardless of the addition of text.
  - **Cropped Images**: To cut out or trim unneeded portions of an image or a page. Image matching from LTU Engine handles cropped images without difficulty.
Composite Images

A composite image contains several photographs or graphics in one image and often has a modified background or added text. For this kind of transformation, LTU Engine's image matching technology delivers extremely accurate results.

3.1.3.2.5. Compression and Image Encoding

In addition to the visually apparent image transformations detailed above, LTU engine is capable of detecting image clones even if the format or compression of the image has changed. Different image file formats include .bmp, .gif, .jpeg, .pov, .png, .psb, .tga, .tif.

Images are often saved in compressed file formats in order to facilitate faster downloading on the Internet. This compression alters the image slightly, but does not typically impact LTU engine's ability to identify a match.

3.1.3.2.6. Images Derived from Mobile Devices

The image matching technology from LTU Engine is optimised to handle query images taken with a mobile device. Due to induced scale changes, motion blur, compression artifacts and usually low quality optics, queries from mobile devices can be challenging to match. LTU Tech has developed an image matching DNA that is particularly robust against combinations of these types of transformations. However, it is recommended to avoid extensive glare, deep angled shots, very dark lighting and to frame the object of interest accordingly.
3.1.3.3. Matching Zone

In addition to visual distance, LTIU Engine is able to return rich information for a query. For example, LTIU Engine can return for each result image, the zones that have matched. This feature is useful:

- to get visual feedback on the algorithm behavior
- to implement custom filtering heuristic (do not return a result if the matching zone is too small) border of uniform color is added on one, several, or all sides of the image.

3.1.3.4. Limitations

3.1.3.4.1. Too important structure modification

The examples below present challenges when matching, due to strong cropping with little structure or too advanced composite images.
3.1.3.4.2. Repetitive pattern

Because repetitive patterns are really similar, pictures parts could be confused or badly identified.

3.1.3.4.3. False Positives

The rate of the false positives depends on the application the image matching technology is integrated with.

Since image matching is very prone to detecting small common parts in images such as logos, it sometimes can result in false positives as seen below because parts of the images actually match.

Sometimes the image matching algorithm detects the same object or scene, but not the same image. According to traditional image recognition search terminology, these instances would be classified as false positives. However, these types of false positives are desirable when performing very fine similarity searches and when the objective is to match photographs taken of objects – proven especially relevant to mobile applications.

3.1.3.4.4. Indexing Limitations

Sometimes an image may not be indexed. This is due either to an unknown image format or due to missing image information.

- Uniform colored images are rejected.
- Having no distinct image features may be rejected too such as the image below.

Finally, images with dimensions of less than 64×64 pixels are rejected in the default value of the LTU Engine (the default setting can be changed).
3.1.4. Similarity Search

3.1.4.1. Visual search

LTU provides a solution for finding similar images. By submitting a query image, our technology can find visually similar images. Similarity can focus on the shapes within the image, its colour or both to:

- recommend similar products for e-Commerce
- navigate through a catalog of images
- return many results, useful for investigation cases

It analyses two characteristics: shapes and colours. These parts are independent and their scores are only merged at the end into the final score of the signature.

- **Shape:**
  Shape similarity is very powerful and can find images regarding different levels of similarity. The algorithm can find images with overall similar shapes. It means that if the query image looks like a ball, we will be able to retrieve other images whose overall shape is a ball as well.

- **Texture:**
  On a thinner level, the algorithm is able to detect the type of texture used in the image. As a result, it finds paintings from the same painter to be similar, if the painter used the same texture techniques on different paintings.

- **Colour:**
  The colour part of the signature is invariant to scale, rotation or any linear transformation. Colour search is quite flexible and can find images sharing the same colors. It also takes proportion of colors into account.

The relative importance of the colour can be set at each query with a colour weight.

- **Colour Weight 0:**
  If the colour weight is zero, then the algorithm will only focus on the similar shapes.

- **Colour Weight 100:**
  With a colour weight at 100, the algorithm will only take colours into account when looking for similar images.

- **Colour Weight 50:**
  An intermediate value between zero and one hundred indicates that both shapes and colours should be taken into account.

3.1.4.2. Search by keywords

LTU solution also allows to search in an images collection by using one or a combination of several keywords. The similarity is not just visual anymore, it is thematic. For example, that could help to find all the pictures of a photograph whose theme is "fruits".

3.1.5. Color Search

3.1.5.1. Overview

Additionally to image matching, LTU Engine provides LTU Color Query.
LTU Engine Colour Query is a powerful tool that analyses the colours in an image. As explained in the part 1 of this document, that allows to:

- find the most popular colour or colour palette in a collection of images
- identify all colours in an image or collection of images:
  - value and percentage

But the colour can also be a criteria of search:

- search for images by colour(s) with optional color ponderation (e.g. 25% red, 75% green)
- upload an image to find images with similar colours

Whereas lots of existing colour tools that require human annotation of the image collection, LTU Engine Colour Query is able to analyse the content of your images and automatically identify the present colours. As the process is fully automatic, it is also very accurate. LTU Engine analyses the colors that are actually present in the images, not only a rough hue. This accuracy allows to look for very specific colour tints in an image collection.

### 3.1.5.2. Uniform Background removal

By default, the signature is computed on the whole image. On some specific cases, this behaviour can be problematic. For instance, in e-commerce the articles are often shown on a uniform background. Thus the algorithm considers the background colour as the article main colour. To tackle this issue, LTU Engine introduces a background removal algorithm that identifies uniform backgrounds and computes the signature only on the foreground image. If no uniform background is detected the signature is computed on the whole image.

### 3.1.5.3. Queries

Once LTU Engine has indexed an image collection, it is possible to run queries on it. There are four kind of queries: get image colours, query by colour, query by image, compute palette.

#### 3.1.5.3.1. Query by colour

With LTU Engine you can search in an image collection using a set of colours. For example, LTU Engine lets you run a query by color like “pink” or “pink and green”. Then LTU Engine returns you a list of images that have the desired colour(s). This list is sorted by relevance. LTU algorithm is very accurate. It is able to look very-specific tints. It is also very robust. The algorithm returns the images with the required color tints at top positions but it also returns images with slightly different tints at higher positions (or at top positions if none of the image contains the required colour tints).

Results for a query by colour “red”:

![Results for red query](image1)

Results for a query by colour “yellow and blue”:

![Results for yellow and blue query](image2)
With LTU Engine it is also possible to specify the desired colour proportion. For instance you can run like a query like ‘look for images with 50% red and 25% yellow’.

Results for a query by colour in varying proportions:

3.1.5.3.2. Query by image

Once you have found an interesting photo (using a query by colour for example) you may want to find similar photos in the collection. That is what query by image is for.

Given an input image, query by image looks for images in your collection that have similar colours.

This feature is useful when:
- there are too many colours in an image to type them all
- you do not know a specific colour code

3.1.6. Interaction with keywords

As explained in the part 1, keywords can be assigned to each image in a collection. Keywords can be attributed manually or be the result of a Deep Learning process. So, keywords can then be used with a visual search process to restrict the query result to some specific categories. For instance it is possible to run a query with the keyword “sofa”. Keywords are compatible with Query by colour and Query by image.

Results of a query by color “red” with keyword “sofa”.

3.2. Link images with Metadata

LTU offers you to associate metadata to your images. So, once an image is recognized you can access all the data you have stored with. Metadata are saved in a separated database.
3.3. OCR

LTU could use OCR algorithm to improve a result given by an image search.

3.4. Fine Comparison of Images

Fine image comparison is a specialised technology especially pertinent to media intelligence applications such as advertising identification.

The Fine Image Comparison process generates these elements:

- **score**: a score is generated which quantifies the visual distance between the two images
- **visual indicators**: two analytical images are generated for each fine comparison effected. These analytical images indicate the zones within the images in which there are variations.

The examples below are typical of the types of images to which Fine Image Comparison is applied:

These two images are identical, except for the pricing details in the lower part of the image. The whitened zones in the image at right indicate the zones in which differences are detected.

The differences between these two images are highlighted in the upper left corner.

The Fine Image Comparison is also a good tool to compare original and counterfeited packagings.

- Unidentified advertisements are compared with a database of known advertisements.
- Certain ads are identified as definite matches.
- Other ads are identified as possible matches, but which need validation (their matching scores may indicate the possibility of variations)
- Fine Image Comparison is applied to pairs of possible matches. The score generated by the Fine Image Comparison determines whether the possible matches should be classified as definite matches or should be examined in a human validation process.