ADDEPTO WHITE PAPER 2020

COMPUTER VISION APPLICATIONS

Table of contents:

addepto

WHAT IS COMPUTER VISION?

Computer vision (CV) is an artificial intelligence-based technology that allows computers to observe the world. By analyzing visual data, this innovation can almost perfectly understand a particular situation, and without missing any factors, find the best solutions or the most reasonable decisions.

The algorithms which stand for computer vision reached an amazing level of accuracy in understanding what is happening around. Today's systems are right in 99% of cases – which makes them more accurate than humans.

Computer Vision applications will become crucial in future automation, visual intensive works like RTG luggage inspection, finding criminals with public cameras, or preventing financial fraud using face recognition. This domain will open new areas of development and help to create new industries.

OBJECT DETECTION

Object Detection is a part of Computer Vision which focuses on detecting various objects on photos like cats, dogs, cars, bikes, humans, etc., by extracting features from pixels and applying deep learning to recognize patterns. One of the main areas of Object Detection is face recognition.

3D SCENE RECONSTRUCTION

Algorithms of Computer Vision are able to reconstruct 3D objects from 2D imagery taken from different angles. As an example, we can acquire a city model from images gathered by drones. We may even create a model of the cave based on a movie recorded inside it.

IMAGE AND VIDEO PRE-PROCESSING

Advanced CV with the use of neural networks can perform image transformations not available for traditional image processing algorithms. As an example, we can artificially increase the number of trees or remove them without noticing an artificial change.

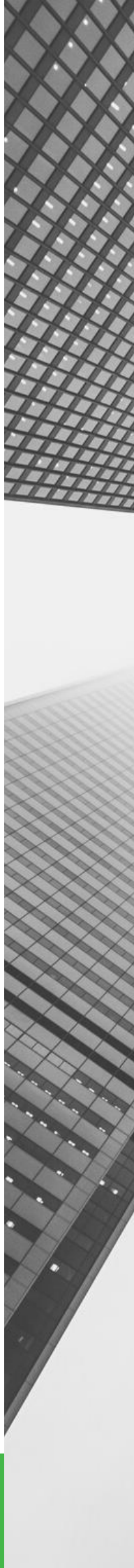
It is possible to generate missing parts of the photo or change the sky's appearance from Earth to Mars. Possibilities of image enhancing and transformation are limitless and require just creating a specialized model for a given task.

SCENE SEGMENTATION

Traditionally, to detect an object on an image it used to be sufficient to just select its position by the rectangle. Now, an improvement of this technique is outlining the given object (for example by a slight change of its color) and in that way segment image on different objects where the result is obtaining an image very similar to the stained glass. This technology will be extensively used in autonomous navigation and radiology (outlining cancerous changes in tissue).

VIDEO AND IMAGE CONTENT INDEXING

A model trained to detect objects on photos can extract its content and prepare tags automatically. Nowadays, the inference is so fast that videos can be processed in real-time. This technology can be used in personalized advertisements (for example screens in public space) where ads are chosen basing on your clothes and things you carry.



WHAT TOOLS ARE USED TO CREATE COMPUTER VISION SOLUTIONS?

C++

C++ is a programming language which supports procedural, object-oriented, and generic programming. It is statically typed, compiled, general-purpose, case-sensitive, free-form framework. It comprises a combination of both high-level and low-level language features.

PYTHON

Python is one of the most popular programming languages in the world. Is being used by companies like Wikipedia, Google, Yahoo!, CERN, and NASA.

It's often used as a "scripting language" for web applications - it can automate a specific series of tasks, making it more efficient. Python is often used in software applications, web pages, and games. It is also used in scientific and mathematical computing, and in AI projects.

OPEN CV

OpenCV library is an open-source computer vision and machine learning software library. It was built primarily to provide an infrastructure for computer vision applications.

OpenCV library has over 2,500 optimized algorithms, which include either the computer vision and machine learning algorithms. These algorithms can be used by companies detect and recognize faces (face recognition), identify objects, classify human actions in videos, track camera movements, track moving objects, extract 3D models of objects, find similar images from an image database, follow eye movements, recognize scenery, and establish markers to overlay it with augmented reality.

COMPUTER VISION TOOLS

TORCH

Torch offers a wide support for machine learning algorithms that puts GPUs (graphics processing units) first. It is very efficient if it comes to fast scripting language, LuaJIT, and an underlying C/CUDA implementation.

PYCHARM

PyCharm is one of many IDEs (integrated development environment) available for Python. It is user-friendly, powerful, and provides integration with git. PyCharm has its own terminal, python console, and provides support for various useful plugins.

KERAS

Keras is a high-level library that uses TensorFlow, CNTK, or Theano as a back-end. It is officially supported by Google (TensorFlow) which has intercepted its development. Keras positions itself as a CV API for "human beings". It focuses on simplicity so creating networks is fast and intuitive.

Model architecture is divided on fully-configurable modules like neural layers, optimizers (Adam, RMSProp), cost functions, etc. It includes built-in models like ResNet50, InceptionV3, or MobileNet. Keras can be used on multi-GPU systems but it requires more time to configure with using both Keras and Tensorflow API.

THEANO

Theano is one of the oldest Python libraries built for operating on multi-dimensional arrays and that allow training neural networks. It is integrated with NumPy, it has efficient symbolic differentiation, possibilities to evaluate expressions faster thanks to dynamic C code generation, and can automatically diagnose many types of errors. Its development has finished in late 2017 but it is still a decent library to use for your project.

COMPUTER VISION TOOLS

TENSORFLOW

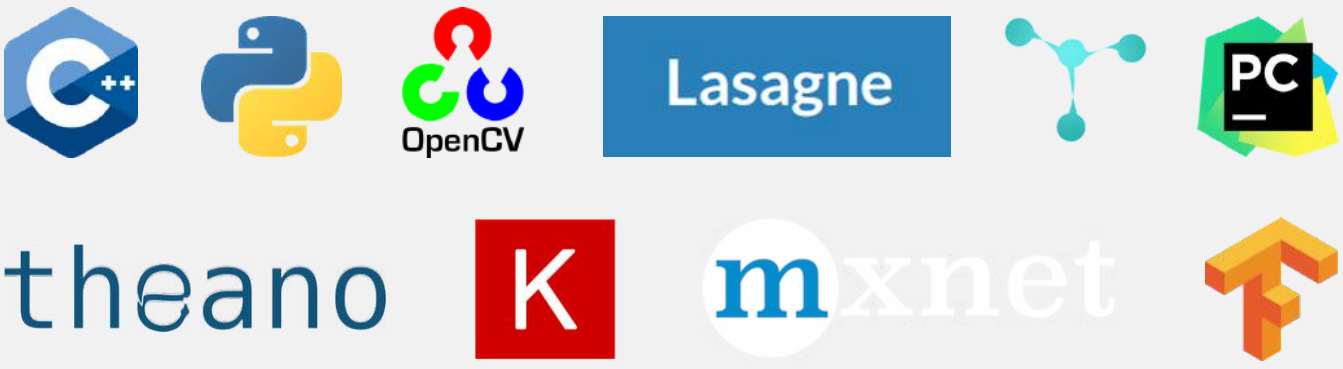
TensorFlow was designed by Google Brain Team and released as an open-source library for abstract (using tensors) numerical computation. It is a low-level library, old enough to have many sophisticated projects using it as a backbone, decent documentation, and vast community. TensorFlow's main advantage (over Theano) is multi-GPU support. It has two API: low-level (original), and high-level Keras.

MXNET

MXNet allows using many GPUs in distributed systems. It is also easy to manage where every piece of data should be stored in the systems. This library has also built-in methods for fast derivative calculations. Every coded layer has been optimized and now MXNet is one of the fastest available CV libraries. However, it takes it more time to start modeling comparing to Keras.

LASAGNE

Lasagne is built on top of Theano with the intention to be simple to understand, use, and easy to directly process and return Theano expression or NumPy data types. Lasagne allows defining Convolutional Neural Networks, Recurrent Neural Networks, and its combinations. It supports CPU and GPU thanks to Theano's compiler. In terms of library level, it is medium – somewhere between low-level libraries like TensorFlow or Theano and high-level libraries like Keras.



REAL-LIFE EXAMPLES OF COMPUTER VISION APPLICATIONS

RETAIL SHELF ANALYSIS

Automatic product detection allows recognizing missing and misplaced products on shelves with comparison to the planogram. Aggregated information about shop conditions gives the opportunity to improve the quality of customer service.

RTG ANALYSIS

Computer Vision can also automate the process of discovering illicit items in luggage during customs inspection on the airports. Such a mundane task is ideal for Convolutional Neural Networks taking into consideration the huge size of the available data-set.

AUTOMATIC VIDEO TAGGING FOR REAL-TIME MARKETING

This technology will improve the advertisement industry, making it more personalized. For example, after tagging customer's favorite brands and gaining deep insights into their preferences, we can recommend products with a higher probability of being chosen. It is a win-win situation for both customers (more relevant ads) and e-commerce (higher income).

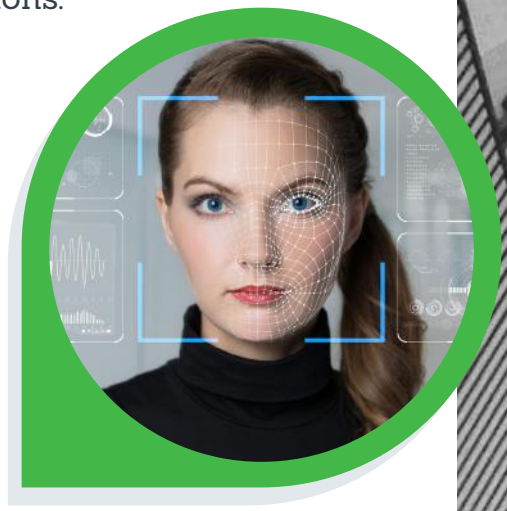
COMPUTER VISION APPLICATIONS

REAL ESTATE VALUATION

Having real estate imagery data with its value, we can create a model that will predict value from new real estate photos. It allows fast comparison of given and predicted prices in order to find investment gems or to find undervalued rent occasions.

RECOGNIZING FACES IN SECURITY SYSTEMS

Make identification easier for security officers and ordinary people – no more need for additional cards or keys. Also, there is a possibility to determine when somebody is a wanted criminal.



AUTOMATIC READING OF PERSONAL INFORMATION FROM IDENTITY CARDS

This technique protects from misspelling and it is much faster than reading information manually. It has the potential to simplify maintaining a customer database and improve the quality of data.

INDUSTRIAL MAINTENANCE

CV techniques use data from cameras to visually check the condition of assets, for example, valves and pipes, and compare it with optimal conditions. This information can be transferred to a remote maintenance crew, that checks anomalies.

ADDEPTO CASE STUDY

BAGGAGE SIMILARITY

CHALLENGE:

There are cases of lost passengers' luggage. Our goal was to find lost luggage at other airports. The problem is that manual search is a very labor intensive task because you have to analyze terabytes of videos.

OUR SOLUTION:

To solve the existing problem, we created the Deep Learning model to find lost luggage. We used FgSegNet for background segmentation, and we used Triplet and Siam networks to find luggage (accuracy in the top five is 94%). The final solution was built in C ++ with an intuitive interface for users. The solutions work in real time using Jetson graphics processors.

BENEFITS:

The prepared solutions processed terabytes of films within a few hours and are looking for lost luggage with great accuracy. It saves many hours of work and optimizes airport costs.

Discover other Addepto case studies.

ABOUT ADDEPTO

Our team builds innovative applications and products by integrating computer vision services with other systems like POS, ERP, and diagnostic software. It is used to detect anomalies in shopping centers, track quality in production lines, analyze medical images, identify products on shelves, and analyze people and their demographics in social media.

AI development experts at Addepto have outstanding experience in building customized computer vision applications with advanced components based on neural networks such as object classification, feature recognition, image segmentation, pattern recognition, object detection, background segmentation, and emotion detection. Those solutions help to solve complex business challenges in different industries.

If you are looking for more details, or you would like to ask us some questions, do not hesitate to contact us anytime.

contact@addepto.com

Visit our website: addepto.com

Find us on social medias: