Applying the Right Deep Learning Model with the Right Data for Your Application

Hila Blecher - Segev
Computer Vision and AI Research Associate, Vision - Elements

2021 embedded VISION summit
VIRTUAL | MAY 25-28

Vision Elements in a Nutshell

Decades of Computer Vision & Artificial Intelligence (AI) accumulated experience

PhDs & MScs in Physics, Mathematics, and Engineering

Industry leaders and disruptive startups among our clients (e.g Medtronic, Philips, Siemens)

Fast & seamless service

Based in Israel & the US

Agenda

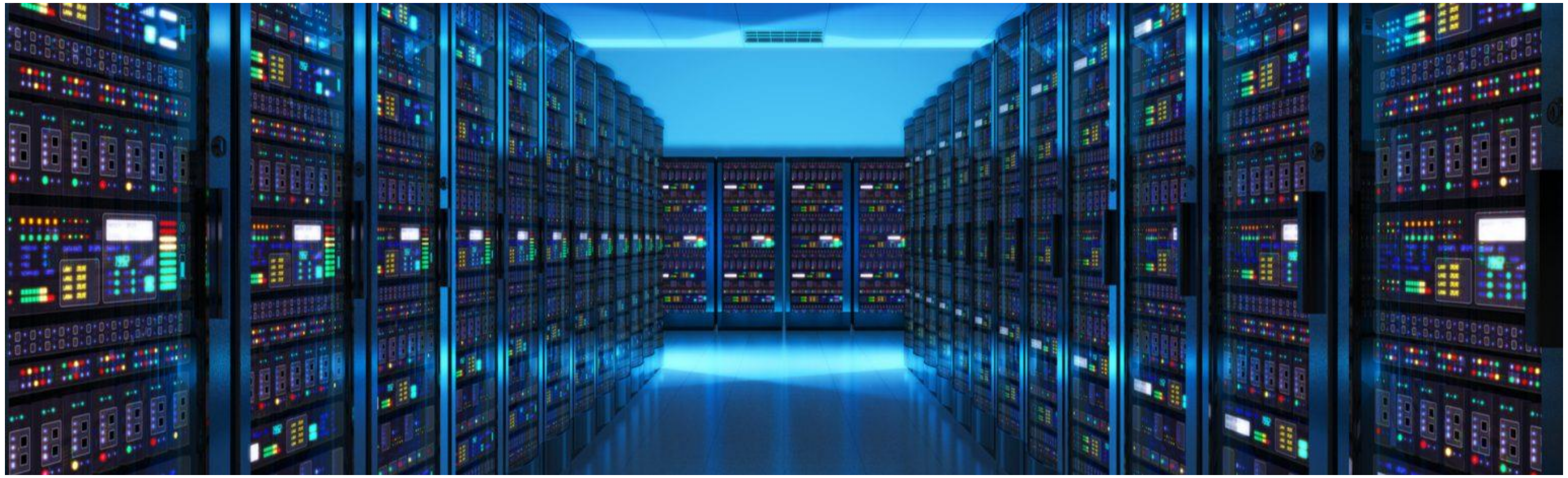
1. Key notes for AI applications
2. Data annotation considerations
3. Algorithmic approaches
4. Additional types of data
5. Model types
6. Volumetric data

1. Key Notes for AI Applications

VisionElements

Key Notes for AI Applications (1)

- **In the past decade we have witnessed the breakthrough of deep neural networks**
- **What is the basis of this black magic?**
  - Endless amount of data with large diversity
  - Improved learning schemes and hardware



Key Notes for AI Applications (2)

- **Data challenges**
  - Large amount of data
    - Small number of annotations fitted for the specific application
  - Data privacy and confidentiality issues
  - Non-balanced data

Semi supervised and unsupervised learning schemes

Transfer Learning

Legalization initiatives

Artificial data creation

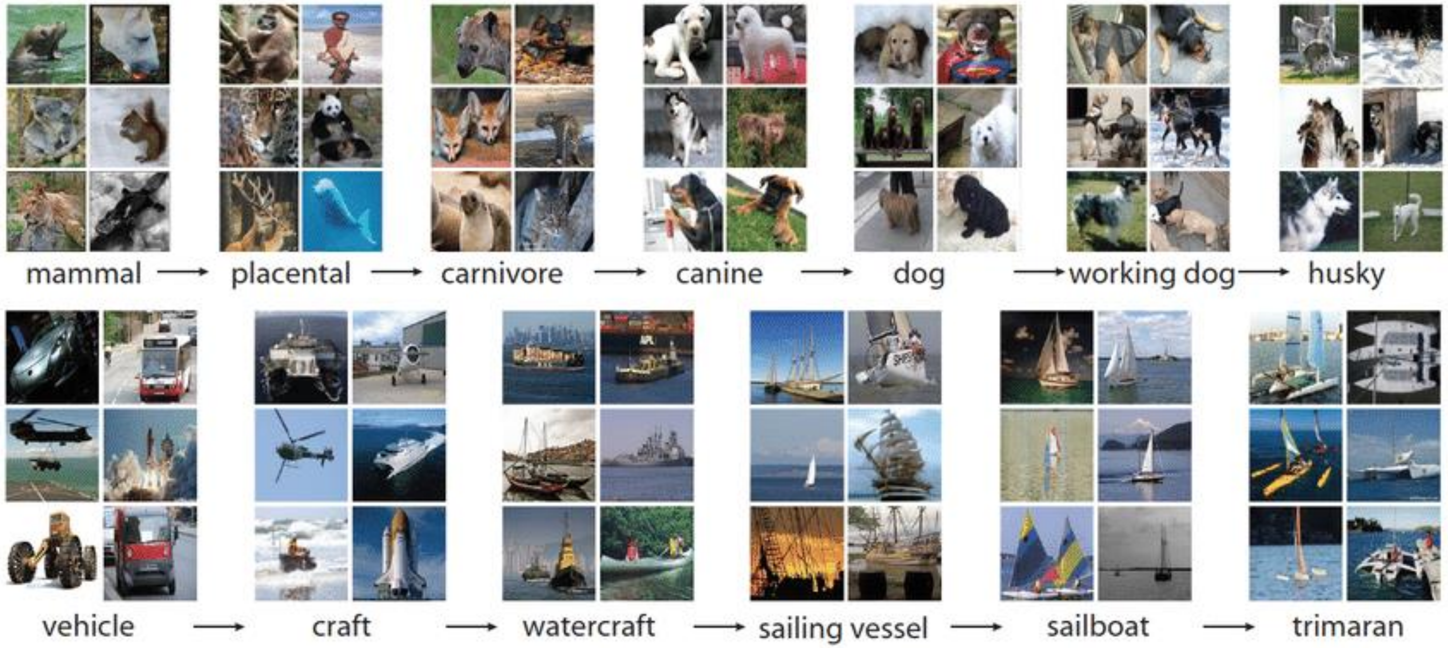
Weighted loss and training

2. Data Annotation Considerations

Data Annotation Considerations (1)

Getting the right data for an application

- Similar characteristic to the data that will be used by the application
  - Indoor\ outdoor
  - Camera angles
  - Lighting conditions
  - Objects
- Large diversity
- Balanced
- Image resolution
- As much data as one can get



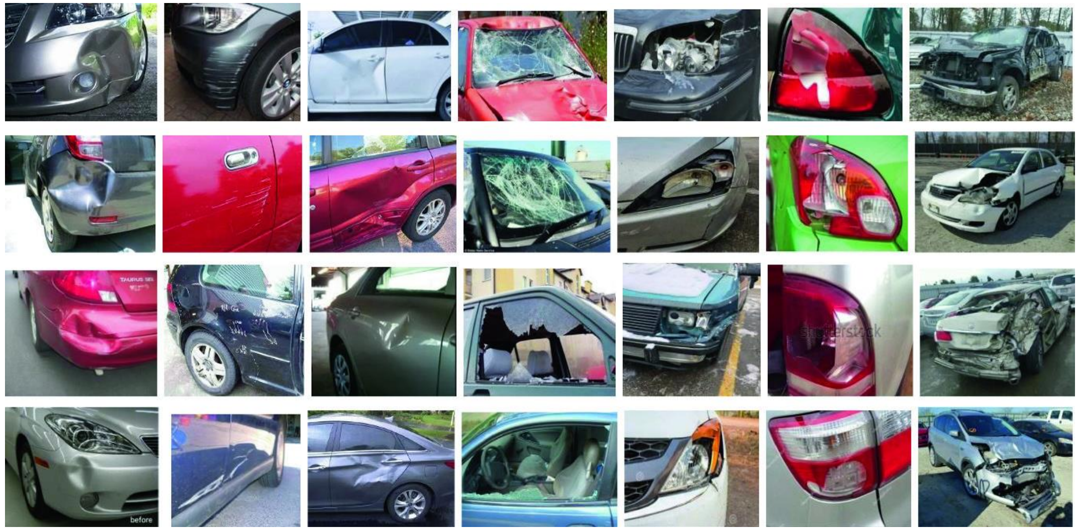
A snapshot of two root-to-leaf branches of ImageNet: mammal sub-tree and vehicle sub-tree. Source: Ye 2018, fig. A1-A

VisionElements

Data Annotation Considerations (2)

Data annotation has a great impact on the model performance

- Prone to human errors
- Requires experts in some cases
- Laborious task



Deep Learning-Based Car Damage Classification and Detection, Mahavir Dwivedi

Drones using AI and Big Data can detect diseased crops from the air, Manchester Metropolitan

Data Annotation Considerations(3)

Map out a process for making good choices for a specific application

What information should we gain from the model?

- Damage Severity → Different classes labeling
- Cost estimation → Additional data
- Damage localization → Segmentation labeling or Bounding Box

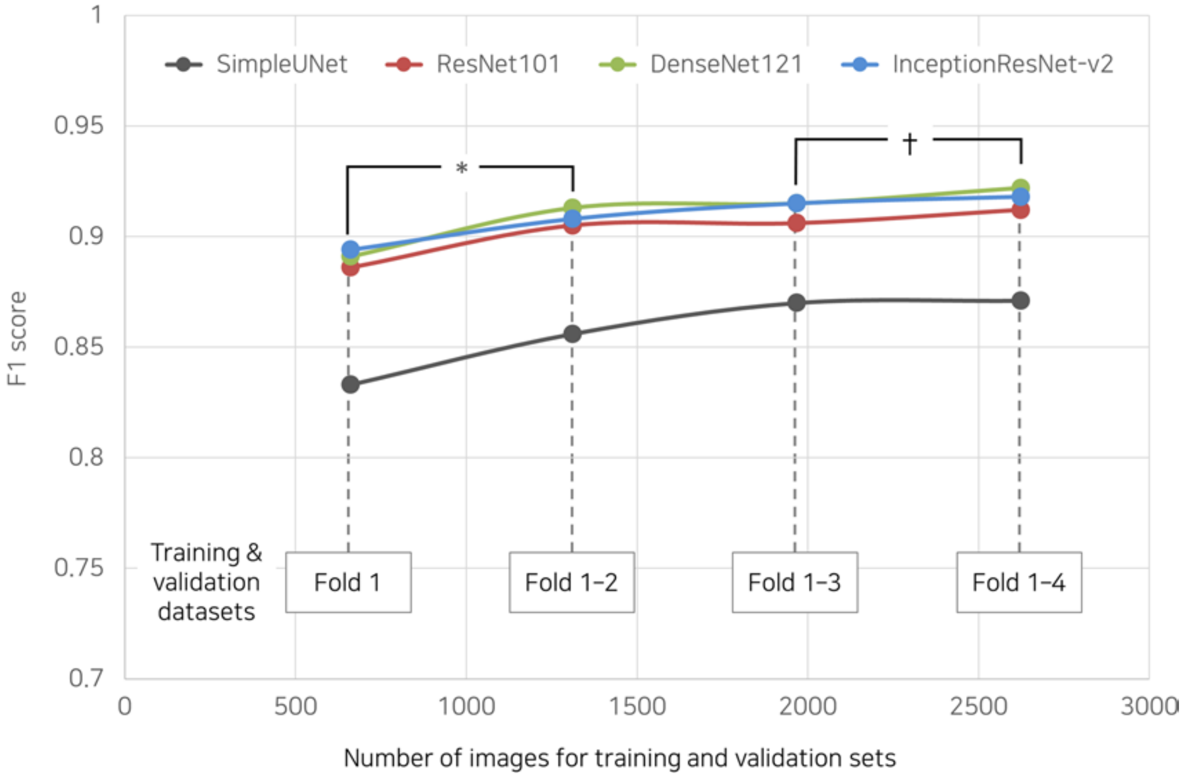
Damage detection from Aerial Imaging

Building Damage Detection from Post-Event Aerial Imagery Using Single Shot Multibox Detector, Yundong Li

Data Annotation Considerations(4)

The million-dollar question: What is the minimum amount of data needed for the network to be generalized and robust?

- Model complexity
- Nature of the problem
- When using transfer learning even small amount of images (thousands) might yield good results



Building Figure 7. Impact of dataset size on segmentation performance of deep learning networks. *p < 0.001 for all deep networks; †p = 0.027 for DenseNet121

3. Algorithmic Approaches

Algorithmic Approaches (1)

Algorithmic approaches to consider depend on the task we are trying to achieve and the data available:

- **Semantic Segmentation:**

  Segmenting different pixels according to their damage type (holes, tarp, exposed plywood, missing shingles)

Pros	Cons
Most detailed information	Time consuming labeling (pixel level)
Good performance of known architectures (Unet, DeepLab, Vnet etc)	Models are more complicated and hence require more data for training
	Annotation is more complicated Boundary areas might be occluded (e.g. trees)

(a) only main buildings (b) with pools

Figure 5: Damage location maps for several examples in D2.

Algorithmic Approaches (2)

- **Classification**

  Classification per deviation from “normal” roof top: major damage , no damage, minor damage and destroyed

Pros	Cons
Classification labeling is simpler and faster (vs segmentation)	Disagreement between labelers (especially within the minor damage class)
Classification architectures are simpler (vs segmentation)	Subjective labeling
Require less data	Less informative
	Classification per whole image VS per house

Algorithmic Approaches (3)

- **Regression object detection**
  - Bounding box location (damaged area localization)
  - Classification to damage types

Pros	Cons
Classification labeling is simpler and faster (vs segmentation)	Disagreement between labelers (especially within the minor damage class)
Bounding box area simpler and faster to mark (vs segmentation)	
Classification architectures are simpler (vs segmentation)	Subjective labeling
Require less data	Less informative

Building Damage Detection from Post-Event Aerial Imagery Using Single Shot Multibox Detector, Yundong Li

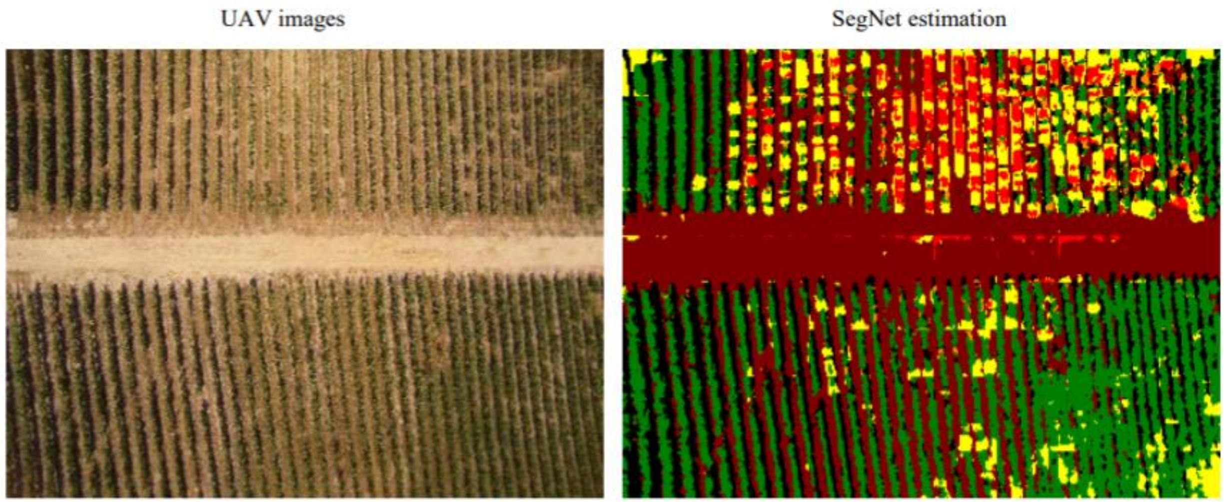
VisionElements

4. Additional Types of Data

Additional Types of Data

Disease detection in multi spectral imaging (Vineyard)

- Type of soil and Previous crop details
- Climate information
- Pre-disease image
- Multiple images from different views or time
- How to add more data?
  - Another input channel
  - Combine in the features (latent) space



Segmentation qualitative results by the SegNet method for the P1 vineyard. On the left, images from UAV and their right, the segmentation result of these images. The color code of the segmentation is; Black: Shadow, Brown: Ground, Green: Healthy, Yellow: Visible symptom, Orange: Infrared symptom, Red: Symptom intersection.

5. Model Types

Model Types (1)

Convolutional Neural Networks (CNN) –
Classification\Regression application

- Mainly used in vision applications, to use image spatial information
- Extract local features from the image
- Fewer parameters than a fully connected layer
- Interleave convolutions, nonlinearities, and (often) pooling operations
- Widely used architectures : Dense net, VGG, Res net

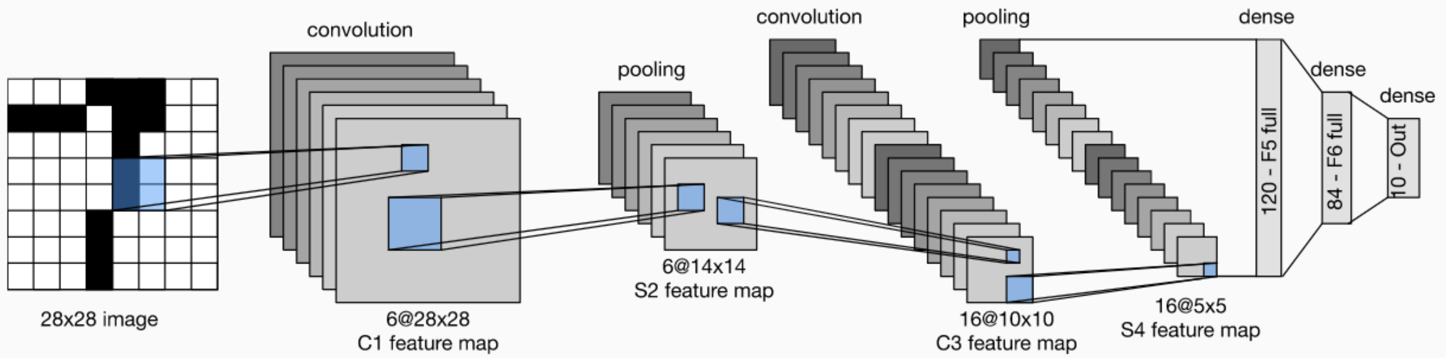
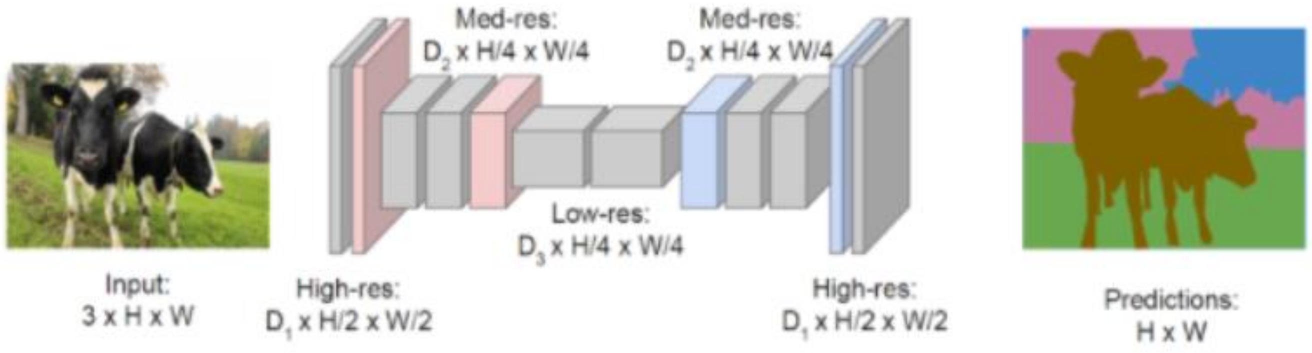


Fig. 6.6.1 Data flow in LeNet. The input is a handwritten digit, the output a probability over 10 possible outcomes.

Model Types (2)

Fully convolutional networks (FCN) - Segmentation

- Uses a CNN to extract important features from the image (encoder)
- Transpose convolutional layers to achieve same spatial resolution at the output (decoder)
- Widely used architectures: Unet, Vnet, Deeplab



Model Types (3)

Generative Adversarial Networks (GAN)

- First component is a generative model that outputs similar data as the real data
- Second component is the discriminator network. It attempts to distinguish real data from fake one
- Both networks are in competition with each other
- Mainly used in anomaly detection, image generation and image translation
- Known architectures: CycleGAN, StyleGAN and text-2image

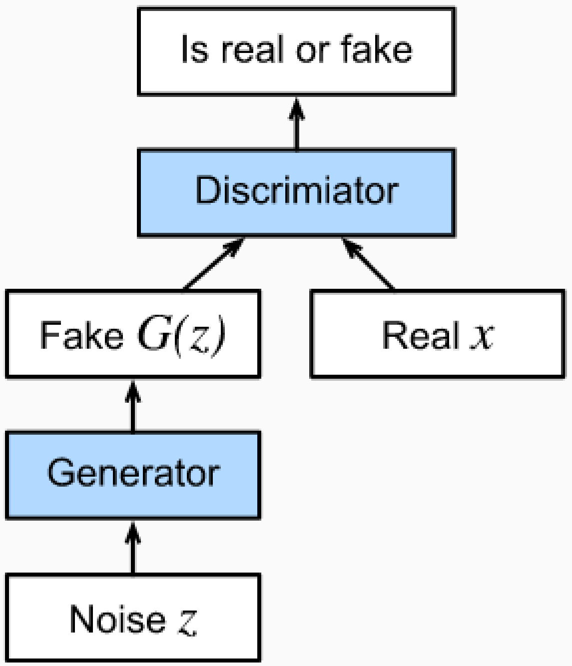


Fig. 17.1.1 Generative Adversarial Networks

6. Volumetric Data

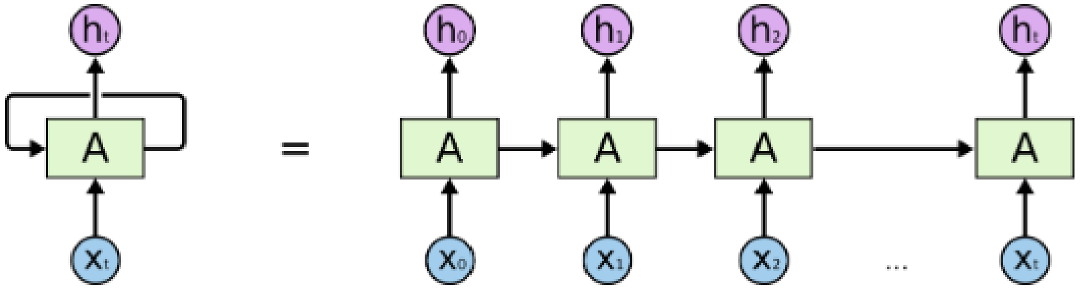
Volumetric Data (1)

- **What would be the best approach for Human Activity Recognition ?**
  - Can we exploit the temporal connection ?

VisionElements

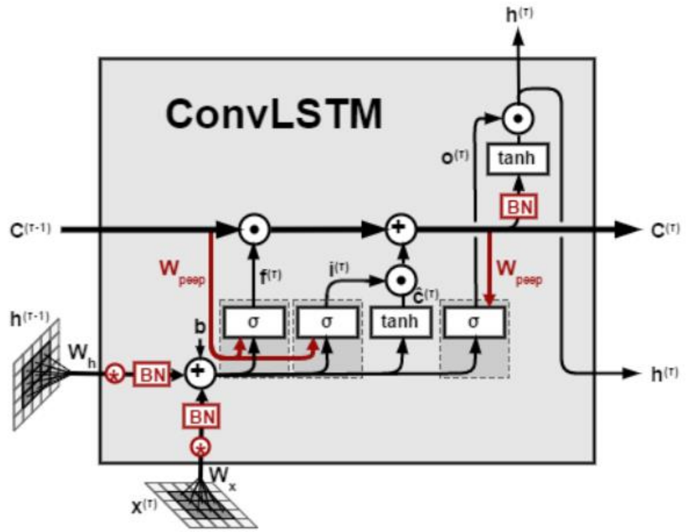
Volumetric Data (2)

Recurrent Neural Network – RNN



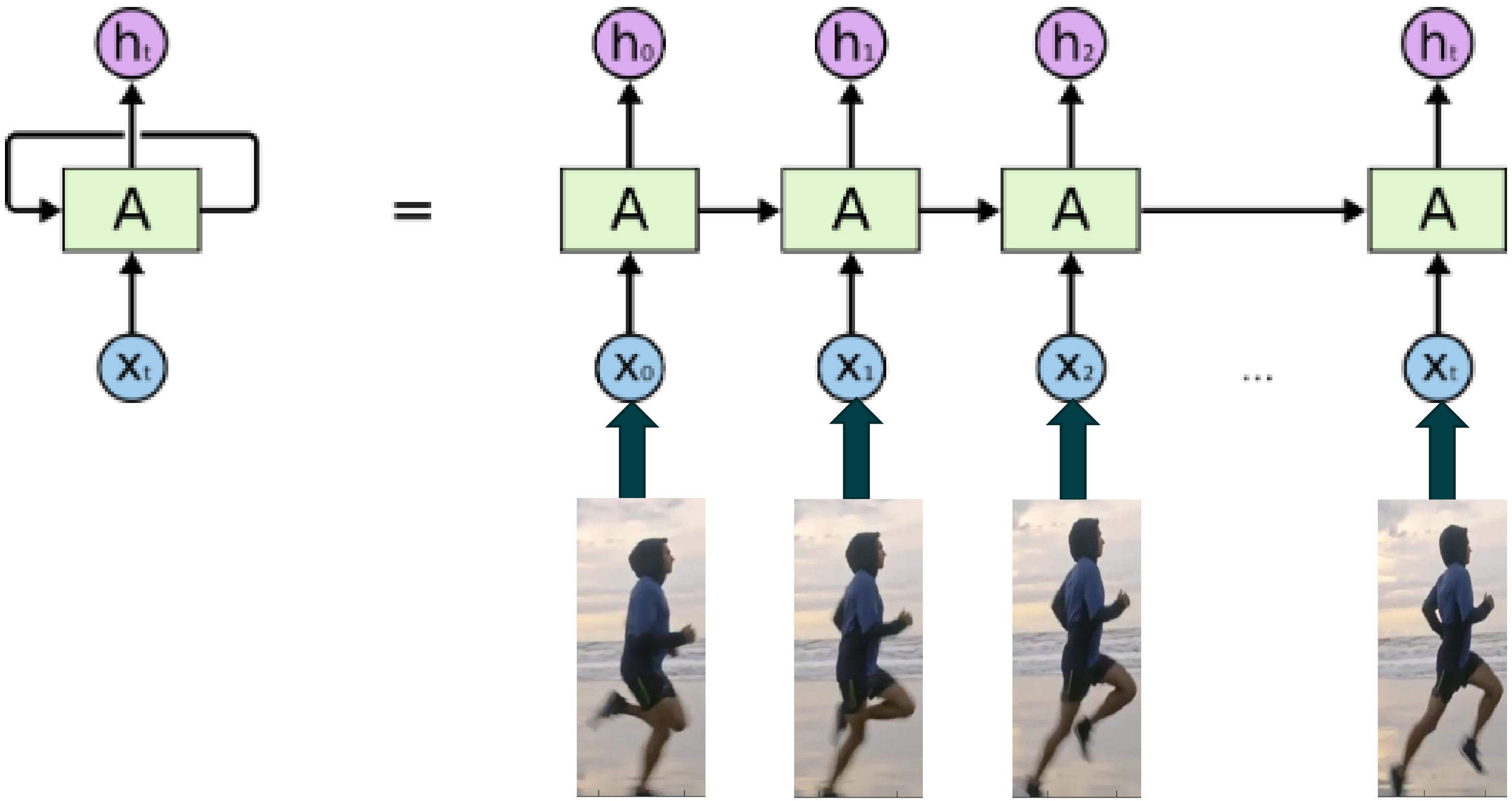
An unrolled recurrent neural network.

Convolution Long Short Term Memory - ConvLSTM



A ConvLSTM cell

Video Incorporation



Key Takeaways

Throughout the presentation we have mapped out the process of making good choices for a specific AI application

- Goal definition
- Data types
- Annotation forms and pitfalls
- Algorithmic approaches and models

These choices are dependent on each other and have a great impact on the final model performance

Resources (1)

- Deep Learning-Based Car Damage Classification and Detection, Mahavir Dwivedi
  https://link.springer.com/chapter/10.1007/978-981-15-3514-7_18

- Drones using AI and Big Data can detect diseased crops from the air, Manchester Metropolitan
  https://www.mmu.ac.uk/news-and-events/news/story/8539/
- Building Damage Detection from Post-Event Aerial Imagery Using Single Shot Multibox Detector, Yundong Li *, Wei Hu, Han Dong and Xueyan Zhang
- Superpixel-wise Assessment of Building Damage from Aerial Images, Lukas Lucks, Dimitri Bulatov
- Vine disease detection in UAV multispectral images with deep learning segmentation approach, Mohamed Kerkech Adel Hafiane and Raphael Canals
- Image net A snapshot of two root-to-leaf branches of ImageNet: mammal sub-tree and vehicle sub-tree. Source: Ye 2018, fig. A1-A
- Su Yang, Jihoon Kweon , Deep learning segmentation of major vessels in X-ray coronary angiography, figure 7

Resources (2)

- Single roof top image from a drone, https://www.everypixel.com/image-1425694916607403839
- Le-Net image for CNN taken from: http://d2l.ai/chapter_convolutional-neural-networks/lenet.html
- GANS https://www.allerin.com/blog/5-applications-of-generative-adversarial-networks
- GANS image : http://d2l.ai/chapter_generative-adversarial-networks/gan.html
- FCN image : https://www.mygreatlearning.com/blog/fcn-fully-convolutional-network-semantic-segmentation/

Thank you for listening

For further information, please contact us:

Hila@vision-elements.com

www.vision-elements.com

Backup slides

Vision Elements services

Computer vision – Straight Forward

3D reconstruction

Object recognition

Segmentation

Classification

Registration

Geo location AR/VR

Tracking

Image/video processing Signal processing

Gesture recognition

Parameter estimation

Point cloud manipulation

Deep learning & AI

Data fusion

VisionElements

Management team

Guy Lavi, MSc

MANAGING PARTNER

Founder, CEO and President at CathWorks

M.Sc. in Geodetic Engineering, Technion, IL

20 years in computer vision, out of which 15years in medical imaging.

Past: Philips Healthcare; Israeli Ministry of Defense; P-Cure

Adi Dafni, MSc

PARTNER

M.Sc. Electrical and Computer Engineering, Tel Aviv Uni, IL

13years in computer vision and machine learning.

Past: IBM, Applied Materials

Asaf Shimshovitz, PhD

PARTNER

PhD. In Physics, Bar Ilan Uni, IL

Developed breakthrough algorithms for solving the Schrodinger equations in quantum mechanics, 4 years of experience in computer vision and machine learning

Adi Sheinfeld, PhD

ASSOCIATE

PhD. in Electrical Engineering, Post-doc in Biomedical Engineering at Duke University

13years in computer vision and machine learning

Past: Applied Materials

Hila Blecher-Segev, MSc

ASSOCIATE

MSc and BSc in Engineering with honors from Ben Gurion University

12 years experience of algorithm development – Computer vision & AI

Past: Rafael – Leading defense company, VR/AR start-up

We're Here For You

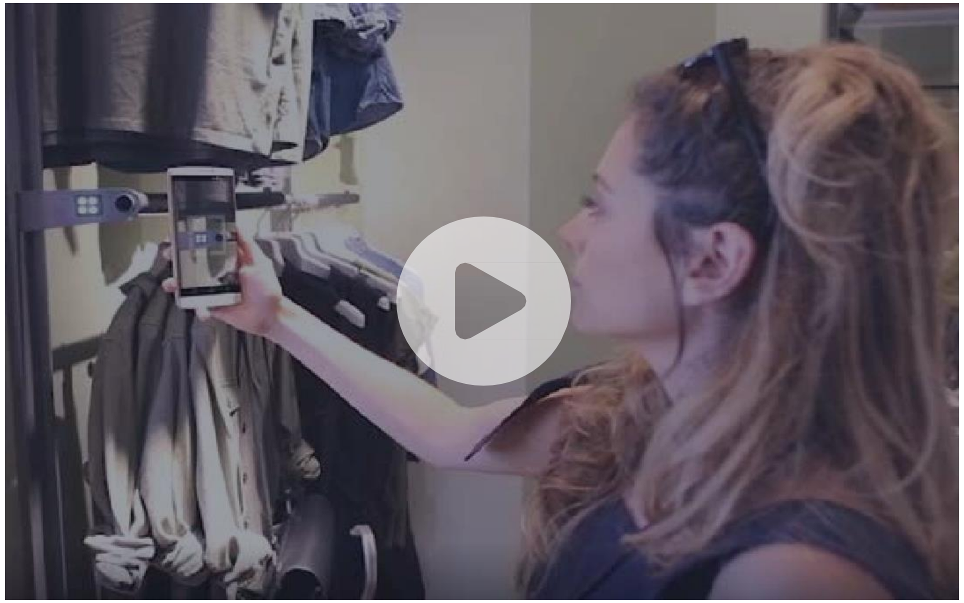
Let us help you with your complex technological obstacles.

- Enjoy rapid deployment
- Increase your team's performance and meet stringent R&D targets
- Utilize a flexible service (short to long term)
- Intellectual property belongs to you
- Stay ahead of the competition

Project example - LED Detection & Color Mapping

Retailers connect physical assets to online content through unique LED-based optical tags.

Vision Elements developed the algorithm – detecting the LEDs and extracting the temporal color vector out of the video.



https://www.vision-elements.com/?video=leds-detection-and-color-mapping

Project example - Gesture Recognition

Multi core vision processor supports 3D imaging.

Vision Elements created depth maps, gesture recognition capabilities and combined near and far depth images into a single HDR image.

https://www.vision-elements.com/?video=hand-gesture-recognition

Project example - Autonomous Drone

Drone follows objects for leisure and sport activities.

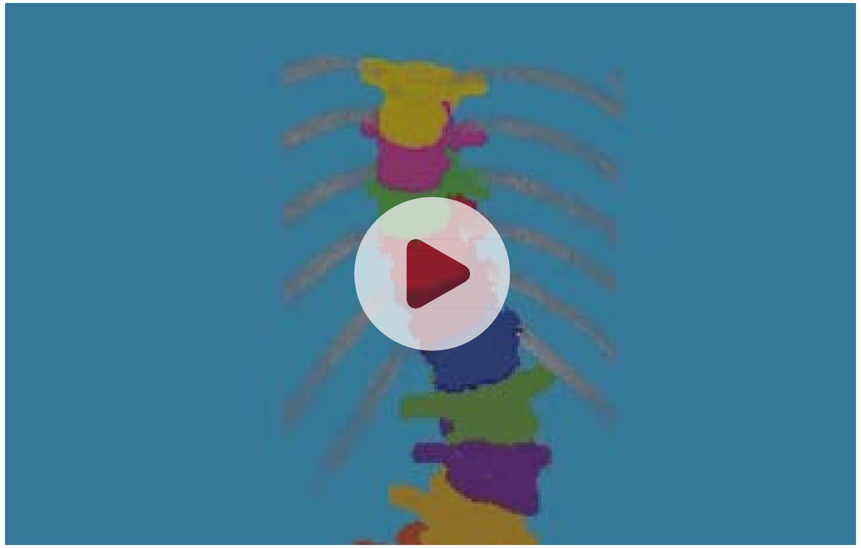
Vision Elements implemented the target tracking – overcoming challenges such as lost view, multiple objects, similar background colors.

https://www.vision-elements.com/?video=autonomous-drone

Project example - Augmented Reality Surgery Planning

CT images with advanced deep learning enable vertebrae labeling and overall surgery planning.

Vision Elements deployed deep learning using 3D convolutional neural network (CNN) for fully automatic segmentation and labeling.



https://www.vision-elements.com/?video=augmented-reality-surgery-planning