# Computer Vision: Challenges and Opportunities

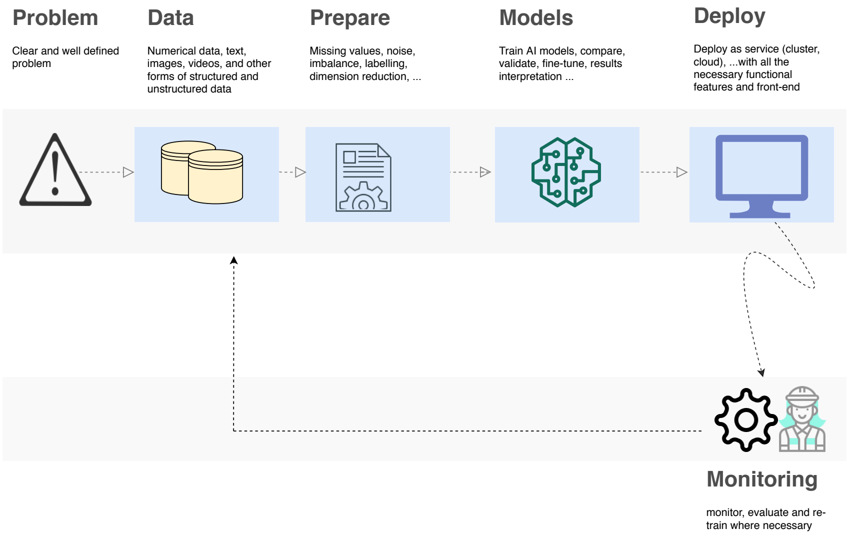

Professor Eyad Elyan

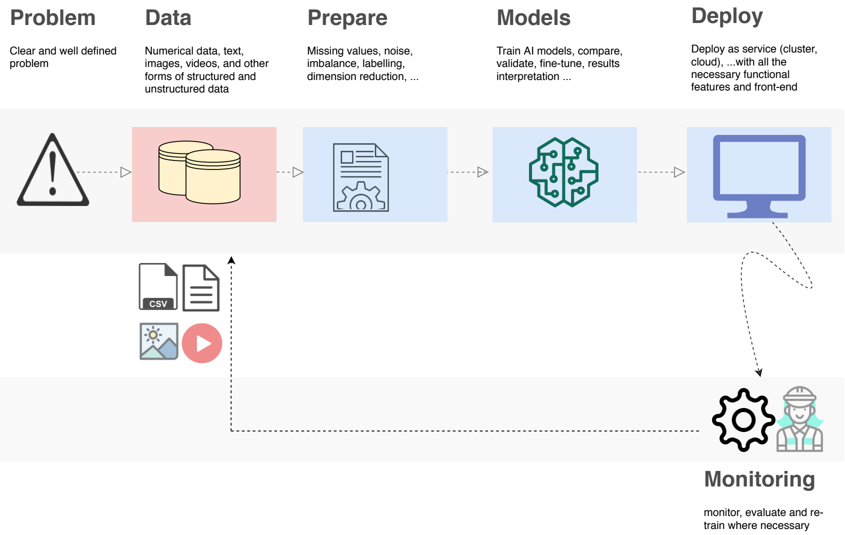
Lead of the Interactive Machine Vision Research Group
School of Computing
Robert Gordon University
Aberdeen, United Kingdom


May 4, 2022

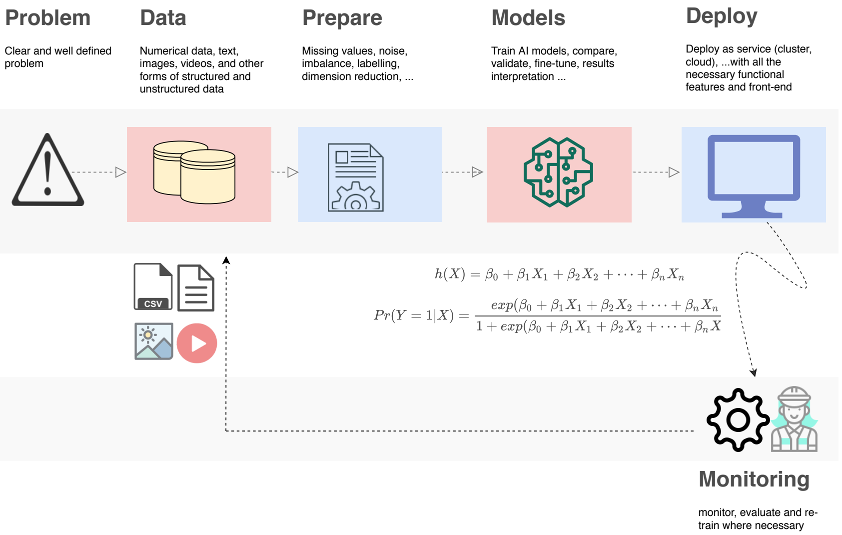
# Overview

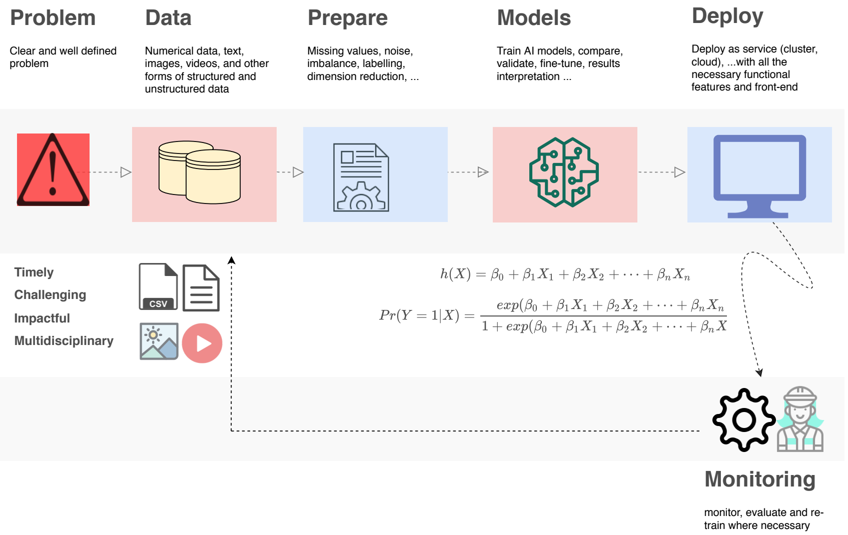
1. Background
    - Computer Vision
    - Image Representation
2. Breakthrough in CV
3. Common CV Tasks
    - Basic Tasks
    - Advanced Tasks
4. Real World Projects
    - Engineering Diagrams
    - Online User's Authentication
    - Remote Inspection
    - Mechanical Engineering Diagrams
5. Challenges
6. Conclusion


Problem
Clear and well defined problem
Data
Numerical data, text, images, videos, and other forms of structured and unstructured data
Prepare
Missing values, noise, imbalance, labelling, dimension reduction, ...
Models
Train AI models, compare, validate, fine-tune, results interpretation ...
Deploy
Deploy as service (cluster, cloud), ...with all the necessary functional features and front-end
Monitoring
monitor, evaluate and re-train where necessary



Problem
Clear and well defined problem
Data
Numerical data, text, images, videos, and other forms of structured and unstructured data
Prepare
Missing values, noise, imbalance, labelling, dimension reduction, ...
Models
Train AI models, compare, validate, fine-tune, results interpretation ...
Deploy
Deploy as service (cluster, cloud), ...with all the necessary functional features and front-end
CSV
Monitoring
monitor, evaluate and re-train where necessary


Problem
Clear and well defined problem
Data
Numerical data, text, images, videos, and other forms of structured and unstructured data
Prepare
Missing values, noise, imbalance, labelling, dimension reduction, ...
Models
Train AI models, compare, validate, fine-tune, results interpretation ...
Deploy
Deploy as service (cluster, cloud), ...with all the necessary functional features and front-end
CSV
$h(X) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n$
$Pr(Y = 1|X) = \frac{exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n}{1 + exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X}$
Monitoring
monitor, evaluate and re-train where necessary


Problem
Clear and well defined problem
Data
Numerical data, text, images, videos, and other forms of structured and unstructured data
Prepare
Missing values, noise, imbalance, labelling, dimension reduction, ...
Models
Train AI models, compare, validate, fine-tune, results interpretation ...
Deploy
Deploy as service (cluster, cloud), ...with all the necessary functional features and front-end
Timely
Challenging
Impactful
Multidisciplinary
CSV
$h(X) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n$
$Pr(Y = 1|X) = \frac{exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n}{1 + exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X}$
Monitoring
monitor, evaluate and re-train where necessary

# Classification - Structured Data

- Predict heart failures in people with cardiovascular disease given a set of input features (e.g. age, sex, type of chest pain, etc...)

| | age | sex | chestPain | bloodpressureRest | serumcholst | fastingbSugar | restingECG | heartRateMax | exercise | oldPeak | slope | NVessels | that | label |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 67 | 0 | 3 | 115 | 564 | 0 | 2 | 160 | 0 | 1.6 | 2 | 0 | 7 | 1 |
| 2 | 57 | 1 | 2 | 124 | 261 | 0 | 0 | 141 | 0 | 0.3 | 1 | 0 | 7 | 2 |
| 3 | 64 | 1 | 4 | 128 | 263 | 0 | 0 | 105 | 1 | 0.2 | 2 | 1 | 7 | 1 |
| 4 | 74 | 0 | 2 | 120 | 269 | 0 | 2 | 121 | 1 | 0.2 | 1 | 1 | 3 | 1 |
| 5 | 65 | 1 | 4 | 120 | 177 | 0 | 0 | 140 | 0 | 0.4 | 1 | 0 | 7 | 1 |
| 6 | 56 | 1 | 3 | 130 | 256 | 1 | 2 | 142 | 1 | 0.6 | 2 | 1 | 6 | 2 |
| 7 | 59 | 1 | 4 | 110 | 239 | 0 | 2 | 142 | 1 | 1.2 | 2 | 1 | 7 | 2 |
| 8 | 60 | 1 | 4 | 140 | 293 | 0 | 2 | 170 | 0 | 1.2 | 2 | 2 | 7 | 2 |
| 9 | 63 | 0 | 4 | 150 | 407 | 0 | 2 | 154 | 0 | 4.0 | 2 | 3 | 7 | 2 |
| 10 | 59 | 1 | 4 | 135 | 234 | 0 | 0 | 161 | 0 | 0.5 | 2 | 0 | 7 | 1 |
| 11 | 53 | 1 | 4 | 142 | 226 | 0 | 2 | 111 | 1 | 0.0 | 1 | 0 | 7 | 1 |
| 12 | 44 | 1 | 3 | 140 | 235 | 0 | 2 | 180 | 0 | 0.0 | 1 | 0 | 3 | 1 |
| 13 | 61 | 1 | 1 | 134 | 234 | 0 | 0 | 145 | 0 | 2.6 | 2 | 2 | 3 | 2 |
| 14 | 57 | 0 | 4 | 128 | 303 | 0 | 2 | 159 | 0 | 0.0 | 1 | 1 | 3 | 1 |
| 15 | 71 | 0 | 4 | 112 | 149 | 0 | 0 | 125 | 0 | 1.6 | 2 | 0 | 3 | 1 |
| 16 | 46 | 1 | 4 | 140 | 311 | 0 | 0 | 120 | 1 | 1.8 | 2 | 2 | 7 | 2 |
| 17 | 53 | 1 | 4 | 140 | 203 | 1 | 2 | 155 | 1 | 3.1 | 3 | 0 | 7 | 2 |
| 18 | 64 | 1 | 1 | 110 | 211 | 0 | 2 | 144 | 1 | 1.8 | 2 | 0 | 3 | 1 |
| 19 | 40 | 1 | 1 | 140 | 199 | 0 | 0 | 178 | 1 | 1.4 | 1 | 0 | 7 | 1 |
| 20 | 67 | 1 | 4 | 120 | 229 | 0 | 2 | 129 | 1 | 2.6 | 2 | 2 | 7 | 2 |
| 21 | 48 | 1 | 2 | 130 | 245 | 0 | 2 | 180 | 0 | 0.2 | 2 | 0 | 3 | 1 |
| 22 | 43 | 1 | 4 | 115 | 303 | 0 | 0 | 181 | 0 | 1.2 | 2 | 0 | 3 | 1 |
| 23 | 47 | 1 | 4 | 112 | 204 | 0 | 0 | 143 | 0 | 0.1 | 1 | 0 | 3 | 1 |
| 24 | 54 | 0 | 2 | 132 | 288 | 1 | 2 | 159 | 1 | 0.0 | 1 | 1 | 3 | 1 |
| 25 | 48 | 0 | 3 | 130 | 275 | 0 | 0 | 139 | 0 | 0.2 | 1 | 0 | 3 | 1 |

# Classification - Structured Data

- Predict heart failures in people with cardiovascular disease given a set of input features (e.g. age, sex, type of chest pain, etc...)

| | age | sex | chestPain | bloodpressureRest | serumcholst | fastingbSugar | restingECG | heartRateMax | exercise | oldPeak | slope | NVessels | that | label |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 67 | 0 | 3 | 115 | 564 | 0 | 2 | 160 | 0 | 1.6 | 2 | 0 | 7 | 1 |
| 2 | 57 | 1 | 2 | 124 | 261 | 0 | 0 | 141 | 0 | 0.3 | 1 | 0 | 7 | 2 |
| 3 | 64 | 1 | 4 | 128 | 263 | 0 | 0 | 105 | 1 | 0.2 | 2 | 1 | 7 | 1 |
| 4 | 74 | 0 | 2 | 120 | 269 | 0 | 2 | 121 | 1 | 0.2 | 1 | 1 | 3 | 1 |
| 5 | 65 | 1 | 4 | 120 | 177 | 0 | 0 | 140 | 0 | 0.4 | 1 | 0 | 7 | 1 |
| 6 | 56 | 1 | 3 | 130 | 256 | 1 | 2 | 142 | 1 | 0.6 | 2 | 1 | 6 | 2 |
| 7 | 59 | 1 | 4 | 110 | 239 | 0 | 2 | 142 | 1 | 1.2 | 2 | 1 | 7 | 2 |
| 8 | 60 | 1 | 4 | 140 | 293 | 0 | 2 | 170 | 0 | 1.2 | 2 | 2 | 7 | 2 |
| 9 | 63 | 0 | 4 | 150 | 407 | 0 | 2 | 154 | 0 | 4.0 | 2 | 3 | 7 | 2 |
| 10 | 59 | 1 | 4 | 135 | 234 | 0 | 0 | 161 | 0 | 0.5 | 2 | 0 | 7 | 1 |
| 11 | 53 | 1 | 4 | 142 | 226 | 0 | 2 | 111 | 1 | 0.0 | 1 | 0 | 7 | 1 |
| 12 | 44 | 1 | 3 | 140 | 235 | 0 | 2 | 180 | 0 | 0.0 | 1 | 0 | 3 | 1 |
| 13 | 61 | 1 | 1 | 134 | 234 | 0 | 0 | 145 | 0 | 2.6 | 2 | 2 | 3 | 2 |
| 14 | 57 | 0 | 4 | 128 | 303 | 0 | 2 | 159 | 0 | 0.0 | 1 | 1 | 3 | 1 |
| 15 | 71 | 0 | 4 | 112 | 149 | 0 | 0 | 125 | 0 | 1.6 | 2 | 0 | 3 | 1 |
| 16 | 46 | 1 | 4 | 140 | 311 | 0 | 0 | 120 | 1 | 1.8 | 2 | 2 | 7 | 2 |
| 17 | 53 | 1 | 4 | 140 | 203 | 1 | 2 | 155 | 1 | 3.1 | 3 | 0 | 7 | 2 |
| 18 | 64 | 1 | 1 | 110 | 211 | 0 | 2 | 144 | 1 | 1.8 | 2 | 0 | 3 | 1 |
| 19 | 40 | 1 | 1 | 140 | 199 | 0 | 0 | 178 | 1 | 1.4 | 1 | 0 | 7 | 1 |
| 20 | 67 | 1 | 4 | 120 | 229 | 0 | 2 | 129 | 1 | 2.6 | 2 | 2 | 7 | 2 |
| 21 | 48 | 1 | 2 | 130 | 245 | 0 | 2 | 180 | 0 | 0.2 | 2 | 0 | 3 | 1 |
| 22 | 43 | 1 | 4 | 115 | 303 | 0 | 0 | 181 | 0 | 1.2 | 2 | 0 | 3 | 1 |
| 23 | 47 | 1 | 4 | 112 | 204 | 0 | 0 | 143 | 0 | 0.1 | 1 | 0 | 3 | 1 |
| 24 | 54 | 0 | 2 | 132 | 288 | 1 | 2 | 159 | 1 | 0.0 | 1 | 1 | 3 | 1 |
| 25 | 48 | 0 | 3 | 130 | 275 | 0 | 0 | 139 | 0 | 0.2 | 1 | 0 | 3 | 1 |

# Classification - Structured Data

- Predict heart failures in people with cardiovascular disease given a set of input features (e.g. age, sex, type of chest pain, etc...)

| | age | sex | chestPain | bloodpressureRest | serumcholst | fastingbSugar | restingECG | heartRateMax | exercise | oldPeak | slope | NVessels | that | label |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 67 | 0 | 3 | 115 | 564 | 0 | 2 | 160 | 0 | 1.6 | 2 | 0 | 7 | 1 |
| 2 | 57 | 1 | 2 | 124 | 261 | 0 | 0 | 141 | 0 | 0.3 | 1 | 0 | 7 | 2 |
| 3 | 64 | 1 | 4 | 128 | 263 | 0 | 0 | 105 | 1 | 0.2 | 2 | 1 | 7 | 1 |
| 4 | 74 | 0 | 2 | 120 | 269 | 0 | 2 | 121 | 1 | 0.2 | 1 | 1 | 3 | 1 |
| 5 | 65 | 1 | 4 | 120 | 177 | 0 | 0 | 140 | 0 | 0.4 | 1 | 0 | 7 | 1 |
| 6 | 56 | 1 | 3 | 130 | 256 | 1 | 2 | 142 | 1 | 0.6 | 2 | 1 | 6 | 2 |
| 7 | 59 | 1 | 4 | 110 | 239 | 0 | 2 | 142 | 1 | 1.2 | 2 | 1 | 7 | 2 |
| 8 | 60 | 1 | 4 | 140 | 293 | 0 | 2 | 170 | 0 | 1.2 | 2 | 2 | 7 | 2 |
| 9 | 63 | 0 | 4 | 150 | 407 | 0 | 2 | 154 | 0 | 4.0 | 2 | 3 | 7 | 2 |
| 10 | 59 | 1 | 4 | 135 | 234 | 0 | 0 | 161 | 0 | 0.5 | 2 | 0 | 7 | 1 |
| 11 | 53 | 1 | 4 | 142 | 226 | 0 | 2 | 111 | 1 | 0.0 | 1 | 0 | 7 | 1 |
| 12 | 44 | 1 | 3 | 140 | 235 | 0 | 2 | 180 | 0 | 0.0 | 1 | 0 | 3 | 1 |
| 13 | 61 | 1 | 1 | 134 | 234 | 0 | 0 | 145 | 0 | 2.6 | 2 | 2 | 3 | 2 |
| 14 | 57 | 0 | 4 | 128 | 303 | 0 | 2 | 159 | 0 | 0.0 | 1 | 1 | 3 | 1 |
| 15 | 71 | 0 | 4 | 112 | 149 | 0 | 0 | 125 | 0 | 1.6 | 2 | 0 | 3 | 1 |
| 16 | 46 | 1 | 4 | 140 | 311 | 0 | 0 | 120 | 1 | 1.8 | 2 | 2 | 7 | 2 |
| 17 | 53 | 1 | 4 | 140 | 203 | 1 | 2 | 155 | 1 | 3.1 | 3 | 0 | 7 | 2 |
| 18 | 64 | 1 | 1 | 110 | 211 | 0 | 2 | 144 | 1 | 1.8 | 2 | 0 | 3 | 1 |
| 19 | 40 | 1 | 1 | 140 | 199 | 0 | 0 | 178 | 1 | 1.4 | 1 | 0 | 7 | 1 |
| 20 | 67 | 1 | 4 | 120 | 229 | 0 | 2 | 129 | 1 | 2.6 | 2 | 2 | 7 | 2 |
| 21 | 48 | 1 | 2 | 130 | 245 | 0 | 2 | 180 | 0 | 0.2 | 2 | 0 | 3 | 1 |
| 22 | 43 | 1 | 4 | 115 | 303 | 0 | 0 | 181 | 0 | 1.2 | 2 | 0 | 3 | 1 |
| 23 | 47 | 1 | 4 | 112 | 204 | 0 | 0 | 143 | 0 | 0.1 | 1 | 0 | 3 | 1 |
| 24 | 54 | 0 | 2 | 132 | 288 | 1 | 2 | 159 | 1 | 0.0 | 1 | 1 | 3 | 1 |
| 25 | 48 | 0 | 3 | 130 | 275 | 0 | 0 | 139 | 0 | 0.2 | 1 | 0 | 3 | 1 |

# Classification - Structured Data

- Predict heart failures in people with cardiovascular disease given a set of input features (e.g. age, sex, type of chest pain, etc...)

| | age | sex | chestPain | bloodpressureRest | serumcholst | fastingbSugar | restingECG | heartRateMax | exercise | oldPeak | slope | NVessels | that | label |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 67 | 0 | 3 | 115 | 564 | 0 | 2 | 160 | 0 | 1.6 | 2 | 0 | 7 | 1 |
| 2 | 57 | 1 | 2 | 124 | 261 | 0 | 0 | 141 | 0 | 0.3 | 1 | 0 | 7 | 2 |
| 3 | 64 | 1 | 4 | 128 | 263 | 0 | 0 | 105 | 1 | 0.2 | 2 | 1 | 7 | 1 |
| 4 | 74 | 0 | 2 | 120 | 269 | 0 | 2 | 121 | 1 | 0.2 | 1 | 1 | 3 | 1 |
| 5 | 65 | 1 | 4 | 120 | 177 | 0 | 0 | 140 | 0 | 0.4 | 1 | 0 | 7 | 1 |
| 6 | 56 | 1 | 3 | 130 | 256 | 1 | 2 | 142 | 1 | 0.6 | 2 | 1 | 6 | 2 |
| 7 | 59 | 1 | 4 | 110 | 239 | 0 | 2 | 142 | 1 | 1.2 | 2 | 1 | 7 | 2 |
| 8 | 60 | 1 | 4 | 140 | 293 | 0 | 2 | 170 | 0 | 1.2 | 2 | 2 | 7 | 2 |
| 9 | 63 | 0 | 4 | 150 | 407 | 0 | 2 | 154 | 0 | 4.0 | 2 | 3 | 7 | 2 |
| 10 | 59 | 1 | 4 | 135 | 234 | 0 | 0 | 161 | 0 | 0.5 | 2 | 0 | 7 | 1 |
| 11 | 53 | 1 | 4 | 142 | 226 | 0 | 2 | 111 | 1 | 0.0 | 1 | 0 | 7 | 1 |
| 12 | 44 | 1 | 3 | 140 | 235 | 0 | 2 | 180 | 0 | 0.0 | 1 | 0 | 3 | 1 |
| 13 | 61 | 1 | 1 | 134 | 234 | 0 | 0 | 145 | 0 | 2.6 | 2 | 2 | 3 | 2 |
| 14 | 57 | 0 | 4 | 128 | 303 | 0 | 2 | 159 | 0 | 0.0 | 1 | 1 | 3 | 1 |
| 15 | 71 | 0 | 4 | 112 | 149 | 0 | 0 | 125 | 0 | 1.6 | 2 | 0 | 3 | 1 |
| 16 | 46 | 1 | 4 | 140 | 311 | 0 | 0 | 120 | 1 | 1.8 | 2 | 2 | 7 | 2 |
| 17 | 53 | 1 | 4 | 140 | 203 | 1 | 2 | 155 | 1 | 3.1 | 3 | 0 | 7 | 2 |
| 18 | 64 | 1 | 1 | 110 | 211 | 0 | 2 | 144 | 1 | 1.8 | 2 | 0 | 3 | 1 |
| 19 | 40 | 1 | 1 | 140 | 199 | 0 | 0 | 178 | 1 | 1.4 | 1 | 0 | 7 | 1 |
| 20 | 67 | 1 | 4 | 120 | 229 | 0 | 2 | 129 | 1 | 2.6 | 2 | 2 | 7 | 2 |
| 21 | 48 | 1 | 2 | 130 | 245 | 0 | 2 | 180 | 0 | 0.2 | 2 | 0 | 3 | 1 |
| 22 | 43 | 1 | 4 | 115 | 303 | 0 | 0 | 181 | 0 | 1.2 | 2 | 0 | 3 | 1 |
| 23 | 47 | 1 | 4 | 112 | 204 | 0 | 0 | 143 | 0 | 0.1 | 1 | 0 | 3 | 1 |
| 24 | 54 | 0 | 2 | 132 | 288 | 1 | 2 | 159 | 1 | 0.0 | 1 | 1 | 3 | 1 |
| 25 | 48 | 0 | 3 | 130 | 275 | 0 | 0 | 139 | 0 | 0.2 | 1 | 0 | 3 | 1 |

$$X = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1p} \\ x_{21} & x_{22} & \dots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{m1} & \vdots & \dots & x_{mp} \end{bmatrix}, Y = \begin{bmatrix} y_1 \\ .. \\ .. \\ .. \\ y_m \end{bmatrix}$$

- **Classification**: If $Y$ is a set of discrete values

## Unstructured Data

Computer Vision and ML: Giving the computer the ability to process, analyse and **synthesise** visual content without explicitly programming it:

- 2D Images
- Videos
- 3D Images
- Multi-modal data (2D &3D, etc..)

- Motivation: achieve faster, more accurate, safer practises, . . .

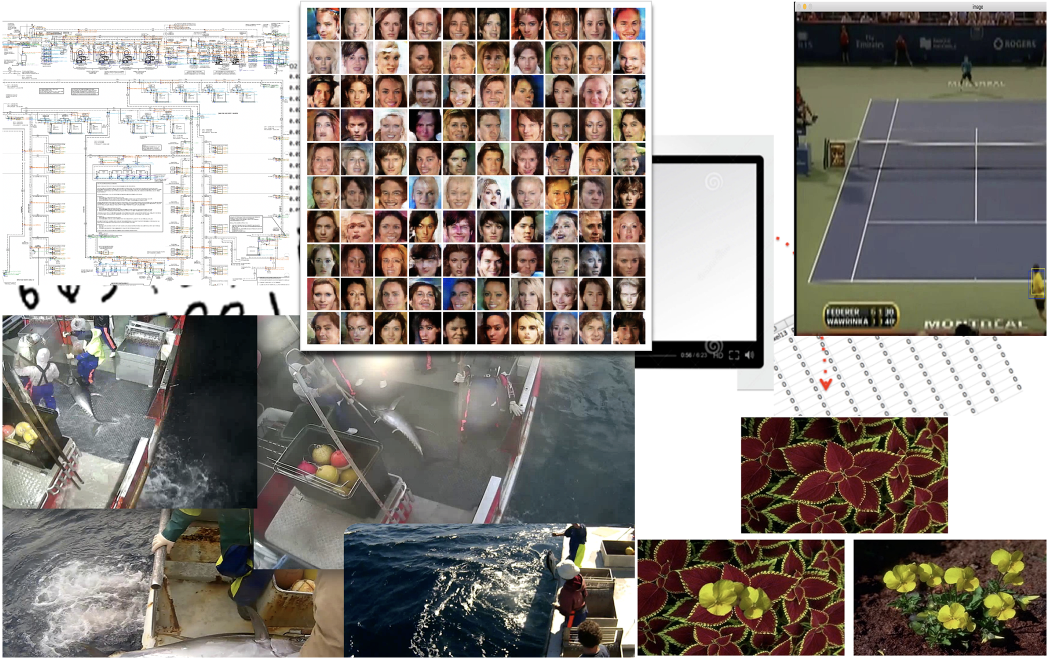


image
ROGERS
FEDERER
WAWRINKA

- *MNIST*: Can computers recognise handwritten digits?

# Image Representation

- *MNIST*: Can computers recognise handwritten digits?

```
| "0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 "
| "0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 "
| "0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 "
| "0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 "
| "0 0 0 0 0 0 0 0 0 1 25 13015525425425415730 2 0 0 0 0 0 0 0 0 "
| "0 0 0 0 0 0 0 0 8 1032532532532532532532531142 0 0 0 0 0 0 0 "
| "0 0 0 0 0 0 0 11 2082532532532532532532532531070 0 0 0 0 0 0 "
| "0 0 0 0 0 0 0 31 2532532532532532532532532532532151013 0 0 0 0 0 "
| "0 0 0 0 0 0 23 21025325325324816122222224625325325325339 0 0 0 0 0 "
| "0 0 0 0 0 0 13625325325322977 0 0 0 70 21825325325325321591 0 0 0 0 "
| "0 0 0 0 0 5 2142532532531950 0 0 0 0 10422425325325325321529 0 0 0 "
| "0 0 0 0 0 11625325325324775 0 0 0 0 0 0 26 2002532532532532164 0 0 "
| "0 0 0 0 0 2542532532531950 0 0 0 0 0 0 0 26 2002532532532535 0 0 "
| "0 0 0 0 0 25425325325399 0 0 0 0 0 0 0 0 0 25 23125325325336 0 0 "
| "0 0 0 0 0 25425325325399 0 0 0 0 0 0 0 0 0 0 2232532532531290 0 "
| "0 0 0 0 0 25425325325399 0 0 0 0 0 0 0 0 0 0 1272532532531290 0 "
| "0 0 0 0 0 25425325325399 0 0 0 0 0 0 0 0 0 0 13925325325390 0 0 "
| "0 0 0 0 0 25425325325399 0 0 0 0 0 0 0 0 0 78 2482532532535 0 0 "
| "0 0 0 0 0 25425325325321634 0 0 0 0 0 0 0 33 1522532532531071 0 0 "
| "0 0 0 0 0 2062532532532531400 0 0 0 0 30 1392342532532531542 0 0 0 "
| "0 0 0 0 0 16 20525325325325020810610610620023725325325325320922 0 0 0 0 "
| "0 0 0 0 0 0 82 25325325325325325325325325325325325325325320922 0 0 0 0 0 "
| "0 0 0 0 0 0 1 91 25325325325325325325325325321390 7 0 0 0 0 0 0 "
| "0 0 0 0 0 0 0 1 18 12920825325325325315912990 4 0 0 0 0 0 0 0 0 "
| "0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 "
| "0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 "
| "0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 "
| "0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 "
```

```
"0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 "
"0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 "
"0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 "
"0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 "
"0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 18825594 0 0 0 0 "
"0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 19125025393 0 0 0 0 "
"0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 12324825316710 0 0 0 0 "
"0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 80 24725320813 0 0 0 0 0 "
"0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 29 20725323577 0 0 0 0 0 0 "
"0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 54 20925325388 0 0 0 0 0 0 0 "
"0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 93 25425323817017 0 0 0 0 0 0 0 "
"0 0 0 0 0 0 0 0 0 0 0 0 0 0 23 2102542531590 0 0 0 0 0 0 0 0 "
"0 0 0 0 0 0 0 0 0 0 0 0 0 16 20925325424081 0 0 0 0 0 0 0 0 0 "
"0 0 0 0 0 0 0 0 0 0 0 0 0 27 25325325413 0 0 0 0 0 0 0 0 0 0 "
"0 0 0 0 0 0 0 0 0 0 0 0 20 2062542541987 0 0 0 0 0 0 0 0 0 0 "
"0 0 0 0 0 0 0 0 0 0 0 0 1682532531967 0 0 0 0 0 0 0 0 0 0 0 "
"0 0 0 0 0 0 0 0 0 0 0 20 20325324876 0 0 0 0 0 0 0 0 0 0 0 0 "
"0 0 0 0 0 0 0 0 0 0 22 18825324593 0 0 0 0 0 0 0 0 0 0 0 0 0 "
"0 0 0 0 0 0 0 0 0 0 1032532531910 0 0 0 0 0 0 0 0 0 0 0 0 0 "
"0 0 0 0 0 0 0 0 0 89 24025319525 0 0 0 0 0 0 0 0 0 0 0 0 0 0 "
"0 0 0 0 0 0 0 0 15 22025325380 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 "
"0 0 0 0 0 0 0 0 94 25325325394 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 "
"0 0 0 0 0 0 0 0 89 2512532501310 0 0 0 0 0 0 0 0 0 0 0 0 0 0 "
"0 0 0 0 0 0 0 0 0 21421895 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 "
"0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 "
"0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 "
"0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 "
"0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 "
```

# Image Representation

```
| "0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  "
| "0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  "
| "0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  "
| "0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  "
| "0  0  0  0  0  0  0  0  0  1  25  130155254254254157 30 2  0  0  0  0  0  0  0  0  0  "
| "0  0  0  0  0  0  0  0  8  103253253253253253253253253 11142  0  0  0  0  0  0  0  "
| "0  0  0  0  0  0  0  11  208253253253253253253253253253253 107 0  0  0  0  0  0  0  "
| "0  0  0  0  0  0  0  31  253253253253253253253253253253253 215 10 13  0  0  0  0  0  "
| "0  0  0  0  0  0  23  210253253253248161222222246253253253253253 39 0  0  0  0  0  "
| "0  0  0  0  0  0  136253253253229 77 0  0  0  70  218253253253253215 91 0  0  0  0  "
| "0  0  0  0  0  5  214253253253195 0  0  0  0  0  104224253253253253215 29 0  0  0  "
| "0  0  0  0  0  116253253253247 75 0  0  0  0  0  0  26  200253253253253216 4  0  0  "
| "0  0  0  0  0  254253253253195 0  0  0  0  0  0  0  26  200253253253 53 5  0  0  "
| "0  0  0  0  0  254253253253 99 0  0  0  0  0  0  0  0  25  231253253253 36 0  0  "
| "0  0  0  0  0  254253253253 99 0  0  0  0  0  0  0  0  0  223253253253129 0  0  "
| "0  0  0  0  0  254253253253 99 0  0  0  0  0  0  0  0  0  127253253253129 0  0  "
| "0  0  0  0  0  254253253253 99 0  0  0  0  0  0  0  0  0  139253253253 90 0  0  "
| "0  0  0  0  0  254253253253 99 0  0  0  0  0  0  0  0  78  248253253253 5  0  0  "
| "0  0  0  0  0  254253253253216 34 0  0  0  0  0  0  33  152253253253107 1  0  0  "
| "0  0  0  0  0  206253253253253140 0  0  0  0  30  139234253253253154 2  0  0  0  "
| "0  0  0  0  0  16  205253253253250208106106106200237253253253253209 22 0  0  0  0  "
| "0  0  0  0  0  0  82  253253253253253253253253253253253253209 22 0  0  0  0  0  "
| "0  0  0  0  0  0  1  91  253253253253253253253253253253213 90 7  0  0  0  0  0  0  "
| "0  0  0  0  0  0  0  1  18  129208253253253253159129 90 4  0  0  0  0  0  0  0  0  "
| "0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  "
| "0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  "
| "0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  "
| "0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  "
```

# Image Representation

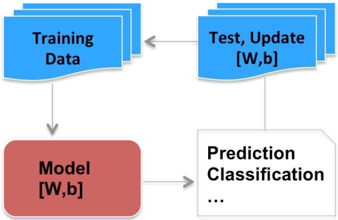
```
| "0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 "
| "0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 "
| "0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 "
| "0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 "
| "0 0 0 0 0 0 0 0 0 1 25 13015525425425415730 2 0 0 0 0 0 0 0 0 "
| "0 0 0 0 0 0 0 0 8 103253253253253253253253253253142 0 0 0 0 0 0 0 "
| "0 0 0 0 0 0 0 11 208253253253253253253253253253253253070 0 0 0 0 0 0 "
| "0 0 0 0 0 0 0 31 253253253253253253253253253253253253253151013 0 0 0 0 0 "
| "0 0 0 0 0 0 23 210253253253248161222222246253253253253339 0 0 0 0 0 "
| "0 0 0 0 0 0 136253253253322977 0 0 0 70 218253253253253321591 0 0 0 0 "
| "0 0 0 0 0 5 214253253253253195 0 0 0 0 0 1042242532532532532321529 0 0 0 "
| "0 0 0 0 0 116253253253324775 0 0 0 0 0 0 26 200253253253253164 0 0 "
| "0 0 0 0 0 254253253253195 0 0 0 0 0 0 0 0 26 200253253253253535 0 0 "
| "0 0 0 0 0 25425325325399 0 0 0 0 0 0 0 0 0 25 231253253253336 0 0 "
| "0 0 0 0 0 25425325325399 0 0 0 0 0 0 0 0 0 0 223253253253131290 0 "
| "0 0 0 0 0 25425325325399 0 0 0 0 0 0 0 0 0 0 127253253253131290 0 "
| "0 0 0 0 0 25425325325399 0 0 0 0 0 0 0 0 0 0 139253253253390 0 0 "
| "0 0 0 0 0 25425325325399 0 0 0 0 0 0 0 0 0 78 248253253253535 0 0 "
| "0 0 0 0 0 254253253253321634 0 0 0 0 0 0 0 33 152253253253131071 0 0 "
| "0 0 0 0 0 206253253253253131400 0 0 0 0 30 139234253253253131542 0 0 0 "
| "0 0 0 0 0 16 205253253253250208106106106200237253253253253320922 0 0 0 0 "
| "0 0 0 0 0 0 82 253253253253253253253253253253253253253253320922 0 0 0 0 0 "
| "0 0 0 0 0 0 1 91 253253253253253253253253253253253321390 7 0 0 0 0 0 0 "
| "0 0 0 0 0 0 0 1 18 129208253253253253159129990 4 0 0 0 0 0 0 0 0 "
| "0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 "
| "0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 "
| "0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 "
| "0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 "
```

| px0 | px1 | px2 | ... | pxn | Label |
|---|---|---|---|---|---|
| 0 | 0 | 0 | ... | 35 | 1 |
| 0 | 0 | 0 | ... | 255 | 3 |
| 0 | 0 | 0 | ... | 0 | 5 |
| 0 | 5 | 0 | ... | 4 | 5 |
| 0 | 3 | 0 | ... | 100 | 9 |
| 0 | 7 | 0 | ... | 10 | 0 |
| 0 | 0 | 0 | ... | 0 | 7 |
| 0 | 0 | 0 | ... | 0 | 7 |
| 0 | 10 | 0 | ... | 0 | 1 |
| 0 | 0 | 0 | ... | 0 | 5 |
| ... | ... | ... | ... | ... | ... |

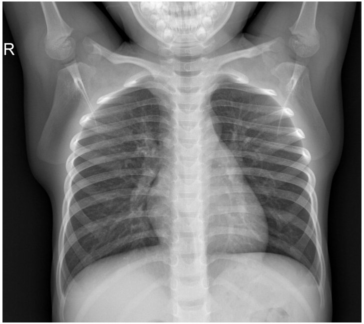
- A 28 $\times$ 28 image would be represented as a 748 row vector in the dataset

# Image Representation for ML

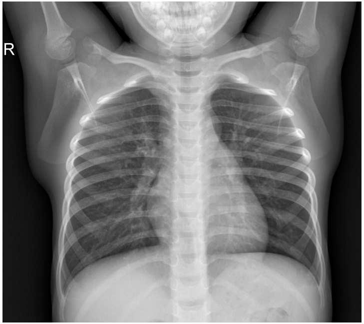
$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & \vdots & \cdots & x_{np} \end{bmatrix}, Y = \begin{bmatrix} y_1 \\ \vdots \\ \vdots \\ y_m \end{bmatrix}$$



# Image Representation for ML



# Image Representation for ML




$$X = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & \vdots & \dots & x_{mn} \end{bmatrix}, Y = \begin{bmatrix} y_1 \\ \cdot\cdot \\ \cdot\cdot \\ \cdot\cdot \\ y_m \end{bmatrix}$$

Each image is represented in a row vector and we want to learn a function h(x) that maps an image $\mathbf{x}_i \in A$ to a class $\mathbf{y}_j \in Y$ e.g. *patient has pneumonia*

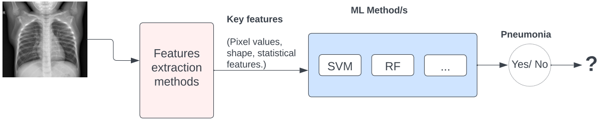
# Plan


1. Background
    - Computer Vision
    - Image Representation
2. Breakthrough in CV
3. Common CV Tasks
    - Basic Tasks
    - Advanced Tasks
4. Real World Projects
    - Engineering Diagrams
    - Online User's Authentication
    - Remote Inspection
    - Mechanical Engineering Diagrams
5. Challenges
6. Conclusion

# Before Deep Learning

Good Old-Fashioned Artificial Intelligence (GOFAI)[1]

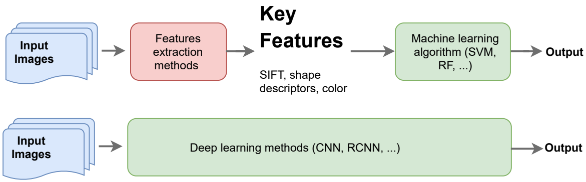
- Hand-craft features (color, shapes, etc...) and heuristics (edge, line detectors, filters, etc...)
- Feed into ML algorithms




[1] Haugeland, John. Artificial intelligence: the very idea. Cambridge, Mass: MIT Press, 1985.

# Deep Learning

- *End-to-End* DL Model that can learn the features and perform the required tasks



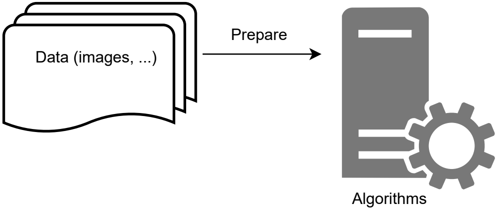

**Paper**

Gumbs, A.A.; Frigerio, I.; Spolverato, G.; Croner, R.; Illanes, A.; Chouillard, E.; Elyan, E. Artificial Intelligence Surgery: How Do We Get to Autonomous Actions in Surgery? Sensors 2021, 21, 5526. https://doi.org/10.3390/s21165526

# Data & Algorithms

Automating complex CV tasks and outperforming humans using DL requires:

- Lots of GOOD quality data (millions of images in some cases)
- Algorithms
- Computing power



# Computing Power

- The cloud
- Super Computers (GPUs)


DGX 1: 1.3 Billion image per day

## Computing Power & Data

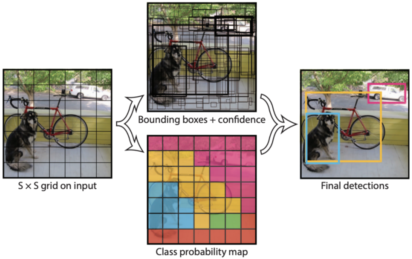
- Computing power (not available)
- Only few 100's of images (or even less)

[2] Darknet code / tutorials https://pjreddie.com/darknet/

# Computing Power & Data

- Computing power (not available)
- Only few 100's of images (or even less)

- With transfer learning you can solve very complex problems (e.g. object detection and tracking using YOLO's darknet implementation[2]




[2] Darknet code / tutorials https://pjreddie.com/darknet/

Significant Progress

## Medical Image Analysis

- Significant progress in image classification, objects recognition and tracking and segmentation [3]

[3] Geert Litjens, Thijs Kooi, Babak Ehteshami Bejnordi, Arnaud Arindra Adiyoso Setio, Francesco Ciompi, Mohsen Ghafoorian, Jeroen A.W.M. van der Laak, Bram van Ginneken, Clara I. Sánchez, A survey on deep learning in medical image analysis, Medical Image Analysis, Volume 42, 2017, Pages 60-88, ISSN 1361-8415

# Medical Image Analysis

- Significant progress in image classification, objects recognition and tracking and segmentation [3]

- Google AI Just Beat Human at Detecting Cancer (89% vs 73% humans accuracy)

[Source: https://www.fool.com/investing/2017/04/04/google-ai-just-beat-human-pathologists-at-detectin.aspx]

[3] Geert Litjens, Thijs Kooi, Babak Ehteshami Bejnordi, Arnaud Arindra Adiyoso Setio, Francesco Ciompi, Mohsen Ghafoorian, Jeroen A.W.M. van der Laak, Bram van Ginneken, Clara I. Sánchez, A survey on deep learning in medical image analysis, Medical Image Analysis, Volume 42, 2017, Pages 60-88, ISSN 1361-8415

# Biometrics (Face Recognition)

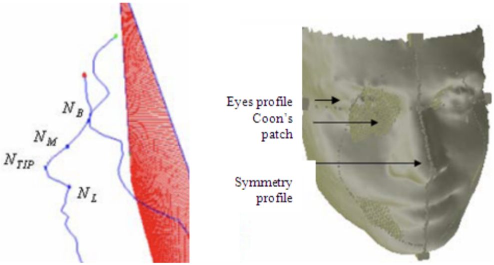
- DeepFace[4] , a face recognition system was first proposed by FaceBook in 2014 achieved an accuracy of 97.35%, beating the state-of-the-art then, by 27%.

- Check code and other interesting CV repositories at https://github.com/ml-tooling/best-of-ml-python

[4] Y. Taigman, M. Yang, M. Ranzato, and L. Wolf, "Deepface: Closing the gap to human-level performance in face verification," in 2014 IEEE Conference on Computer Vision and Pattern Recognition, June 2014, pp. 1701–1708

# Biometrics (Face Recognition)

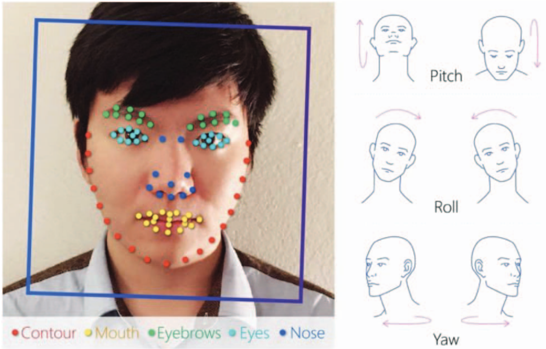
- DeepFace[4] , a face recognition system was first proposed by FaceBook in 2014 achieved an accuracy of 97.35%, beating the state-of-the-art then, by 27%.
- Check code and other interesting CV repositories at https://github.com/ml-tooling/best-of-ml-python




[4] Y. Taigman, M. Yang, M. Ranzato, and L. Wolf, "Deepface: Closing the gap to human-level performance in face verification," in 2014 IEEE Conference on Computer Vision and Pattern Recognition, June 2014, pp. 1701–1708

# Face Analysis and Sexual Orientation

- Using less than 40,000 images, DL outperformed humans in detecting people's sexual orientation from single images[5]




[5] Wang, Y., & Kosinski, M. (2018). Deep neural networks are more accurate than humans at detecting sexual orientation from facial images. Journal of Personality and Social Psychology, 114(2), 246–257. https://doi.org/10.1037/pspa0000098

# Generative Adverserial Neural Network (GAN)

- GANs were developed in 2014 by Ian J Goodfellow

**Paper**

Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, Yoshua Bengio, Generative adversarial nets

# GANs

- GANs have been successfully applied to various applications
- For example face image generation (very convincing)

# GANs

- GANs have been successfully applied to various applications
- For example face image generation (very convincing)

[Source https://thispersondoesnotexist.com/]

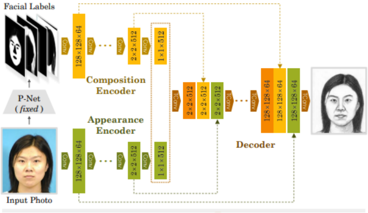
# GANs

- GANs are widely used for generating super-resolution images, image editing and colourisation
- Check *DeOldify*, a powerful GAN-based tool for editing/ coloring images (code is available on github)


[Source https://github.com/jantic/DeOldify]

# GANs

- Wide range of other applications such as generating sketches from images
- Check the GANs Zoo for the list of applications and corresponding source code repository[6] or GANs Awesome Applications[7]




[Yu et al. https://arxiv.org/pdf/1712.00899.pdfTowards Realistic Face Photo-Sketch Synthesis via Composition-Aided GANs]

[6] https://github.com/hindupuravinash/the-gan-zoo
[7] https://github.com/nashory/gans-awesome-applications

# Latest Trending Applications


- Unexpected obstacles detection[8]

- Computer vision in health and safety (e.g. identify potential hazards in construction sites),
- Retail industry[9]

[8] S. Ramos, S. Gehrig, P. Pinggera, U. Franke and C. Rother, "Detecting unexpected obstacles for self-driving cars: Fusing deep learning and geometric modeling," 2017 IEEE Intelligent Vehicles Symposium (IV), 2017, pp. 1025-1032, doi: 10.1109/IVS.2017.7995849.

[9] https://www.forbes.com/

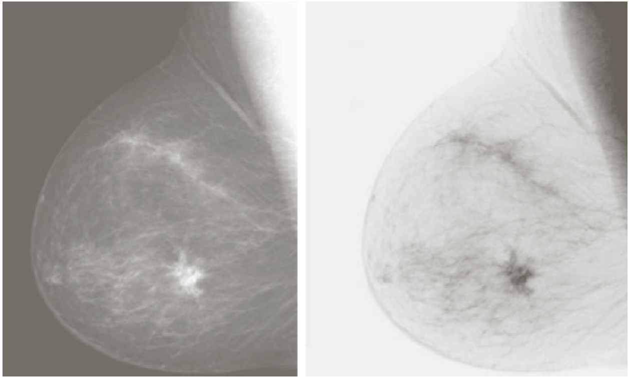
# Plan


1. Background
   - Computer Vision
   - Image Representation
2. Breakthrough in CV
3. Common CV Tasks
   - Basic Tasks
   - Advanced Tasks
4. Real World Projects
   - Engineering Diagrams
   - Online User's Authentication
   - Remote Inspection
   - Mechanical Engineering Diagrams
5. Challenges
6. Conclusion

# Basic Tasks

# Basic CV Methods

- Methods include thresholding, contrast stretching, morphological operations, negation of images, . . . ...

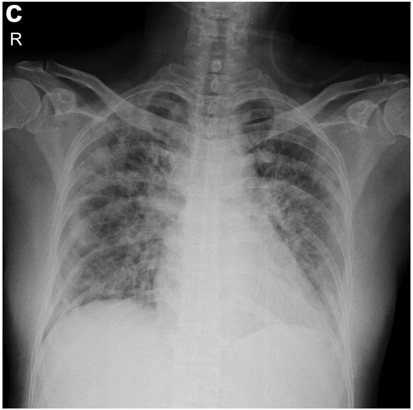


[source: Gonzalez and Woods]

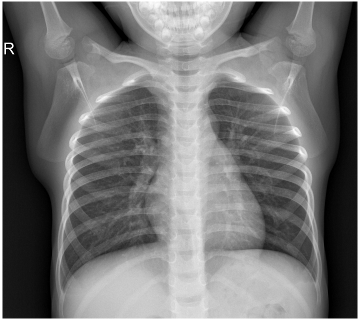
# More Complex Tasks

# Classification

*Covid-19 positive*

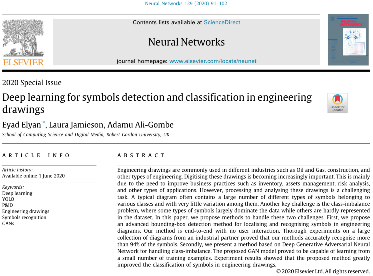



*Covid-19 negative*




[Source: https://github.com/ieee8023/covid-chestxray-dataset]

# Object Detection/ Recognition



Neural Networks 129 (2020) 91–102

Contents lists available at ScienceDirect

Neural Networks

journal homepage: www.elsevier.com/locate/neunet

2020 Special Issue

## Deep learning for symbols detection and classification in engineering drawings

Eyad Elyan *, Laura Jamieson, Adamu Ali-Gombe

School of Computing Science and Digital Media, Robert Gordon University, UK

ARTICLE INFO

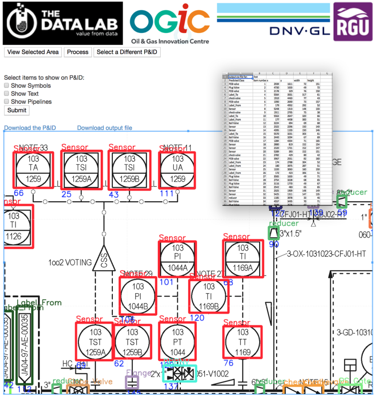
Article history:
Available online 1 June 2020

Keywords:
Deep learning
YOLO
P&ID
Engineering drawings
Symbols recognition
GANs

ABSTRACT

Engineering drawings are commonly used in different industries such as Oil and Gas, construction, and other types of engineering. Digitising these drawings is becoming increasingly important. This is mainly due to the need to improve business practices such as inventory, assets management, risk analysis, and other types of applications. However, processing and analysing these drawings is a challenging task. A typical diagram often contains a large number of different types of symbols belonging to various classes and with very little variation among them. Another key challenge is the class-imbalance problem, where some types of symbols largely dominate the data while others are hardly represented in the dataset. In this paper, we propose methods to handle these two challenges. First, we propose an advanced bounding-box detection method for localising and recognising symbols in engineering diagrams. Our method is end-to-end with no user interaction. Thorough experiments on a large collection of diagrams from an industrial partner proved that our methods accurately recognise more than 94% of the symbols. Secondly, we propose a method based on Deep Generative Adversarial Neural Network for handling class-imbalance. The proposed GAN model proved to be capable of learning from a small number of training examples. Experiment results showed that the proposed method greatly improved the classification of symbols in engineering drawings.



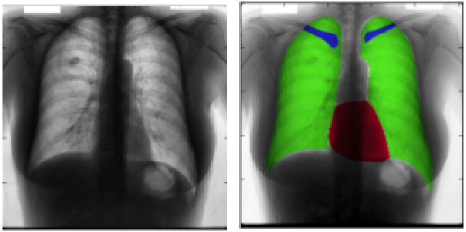



## Paper

E Elyan, L Jamieson, A Ali-Gombe, Deep learning for symbols detection and classification in engineering drawings, Neural Networks,Volume 129, 2020, Pages 91-102, ISSN 0893-6080, https://doi.org/10.1016/j.neunet.2020.05.025

# Segmentation

- Widely used across various applications such as self-driving cars, and very common in biomedical applications



**Paper**

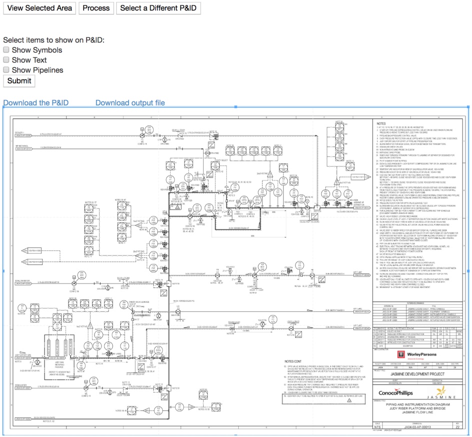
A. A. Novikov, D. Lenis, D. Major, J. Hladůvka, M. Wimmer and K. Bühler, "Fully Convolutional Architectures for Multiclass Segmentation in Chest Radiographs," in IEEE Transactions on Medical Imaging, vol. 37, no. 8, pp. 1865-1876, Aug. 2018, doi: 10.1109/TMI.2018.2806086.

# Plan

1 Background
    Computer Vision
    Image Representation
2 Breakthrough in CV
3 Common CV Tasks
    Basic Tasks
    Advanced Tasks
4 Real World Projects
    Engineering Diagrams
    Online User's Authentication
    Remote Inspection
    Mechanical Engineering Diagrams
5 Challenges
6 Conclusion

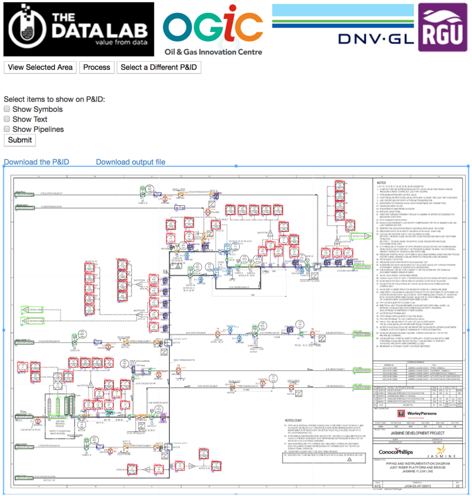
# Engineering Diagrams

- Classify and localise the area of interest



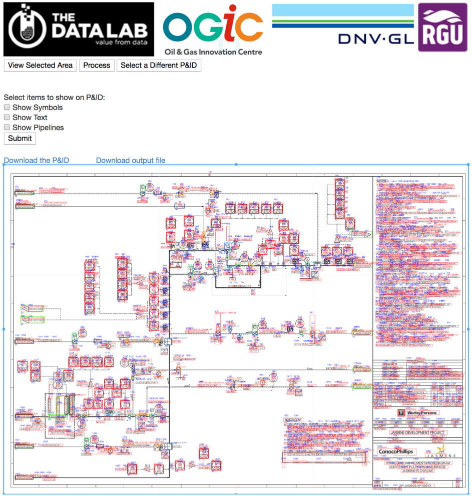
# Engineering Diagrams

- Classify and localise the area of interest



# Engineering Diagrams

- Classify and localise the area of interest



# On-line User's Authentication

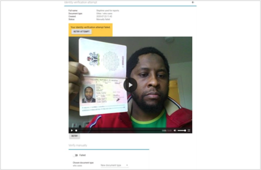
## Research

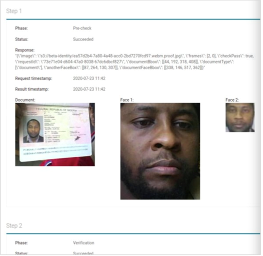
- Deep Convolutional Neural Networks
- Object Detection
- Ensemble Learning

## Application

- Remote user's authentication
- Value $\approx$*£180K*

- *Aug-2019* to *Jul-2021*






Partners- In collaboration with Mintra Group - Aberdeen, UK

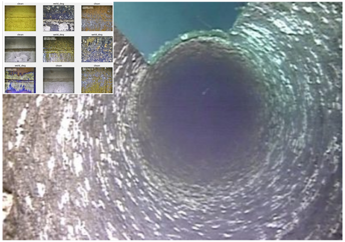
# Remote Inspection

## Research

- Ultrasonic image analysis using Deep Learning-based methods

## Application

- Remote inspection and structural integrity of offshore assets
- KTP with AISUS
- Value $\approx$*£175K*
- *Start date June-2021*



Partners- In collaboration with AISUS LTD - Aberdeen, UK

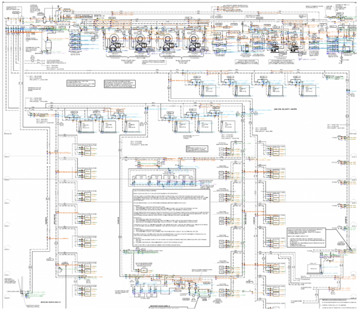
# Mechanical Engineering Diagram Analysis

## Research

- Image analysis
- Deep Learning
- Ensemble Learning

## Application

- Automatic interpretation of mechanical engineering drawings
- Value $\approx$*£300K*
- *Start date Jan-2021*

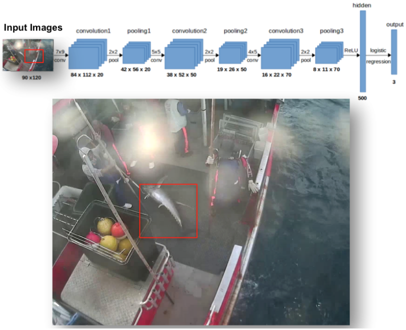


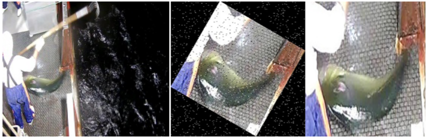
Partners- In collaboration with Fieri Analytics - Canada

# Plan


1. Background
   - Computer Vision
   - Image Representation
2. Breakthrough in CV
3. Common CV Tasks
   - Basic Tasks
   - Advanced Tasks
4. Real World Projects
   - Engineering Diagrams
   - Online User's Authentication
   - Remote Inspection
   - Mechanical Engineering Diagrams
5. Challenges
6. Conclusion

# Data Annotation



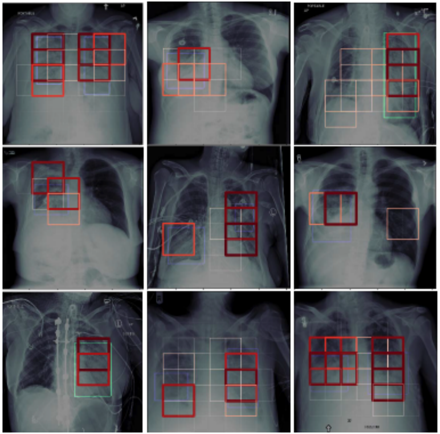



- Better quality and more annotation leads to better results [10]

[10] Adamu Ali-Gombe, Eyad Elyan, Chrisina Jayne, "Fish Classification in Context of Noisy Images". International Conference of Engineering Applications of Neural Networks (EANN) 2017: 216-226,DOI: https://doi.org/10.1007/978-3-319-65172-9_19

# Data Availability/ Annotation

- Availability of good, and accurately annotated datasets of images and videos

 [11]

[11] Source: E. Schwab, A. Gooßen, H. Deshpande and A. Saalbach, "Localization of Critical Findings in Chest X-Ray Without Local Annotations Using Multi-Instance Learning," 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI), 2020, pp. 1879-1882, doi: 10.1109/ISBI45749.2020.9098551.

# Data Generation/ Annotation?

- Generate data using GAN-based models [12]

[12] Adamu Ali-Gombe, Eyad Elyan, MFC-GAN: Class-imbalanced dataset classification using Multiple Fake Class Generative Adversarial Network, Neurocomputing, Volume 361, 2019, Pages 212-221, ISSN 0925-2312,https://doi.org/10.1016/j.neucom.2019.06.043.

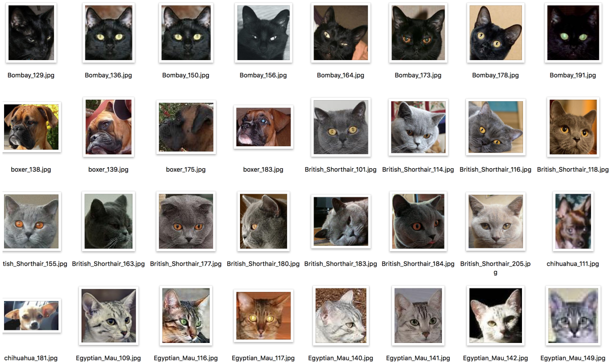
# Data Annotation Tools

Data Annotation is largely manual process using some open source tools:

1. VGG Image Annotator
2. LabelImg
3. OpenLabeler
4. Make Sense https://www.makesense.ai/
5. ImgLab https://imglab.in/
6. Sloth https://github.com/cvhciKIT/sloth
7. ...

# Biased Data

- Models tend to be biased when they learn from imbalanced datasets

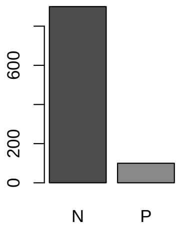


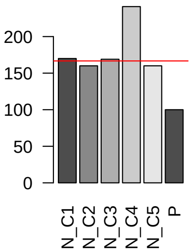

[Source: http://www-edlab.cs.umass.edu/ smaji/cmpsci670/fa14/hw/recognition/]

Kaggle competition https://www.kaggle.com/c/dogs-vs-cats

# Imbalanced Datasets

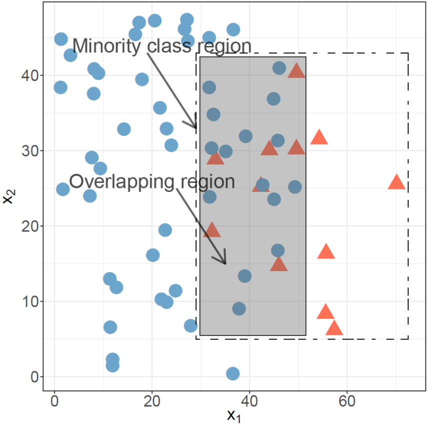
- Find within-class similarity in the dominant class and
- oversample minority-class instances
- Classify and compare[13]




[13] Elyan, E., Moreno-Garcia, C.F. & Jayne, C. CDSMOTE: class decomposition and synthetic minority class oversampling technique for imbalanced-data classification. Neural Comput & Applic 33, 2839–2851 (2021). https://doi.org/10.1007/s00521-020-05130-z

# Imbalanced Datasets





Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins


## Neighbourhood-based undersampling approach for handling imbalanced and overlapped data


Pattaramon Vuttipittayamongkol*, Eyad Elyan

*School of Computing Sciences and Digital Media, Robert Gordon University, UK*



ARTICLE INFO

*Article history:*
Received 9 February 2019
Revised 31 July 2019
Accepted 26 August 2019
Available online 3 September 2019

*Keywords:*
Imbalanced dataset
Undersampling
k-NN
Class overlap
Classification



ABSTRACT

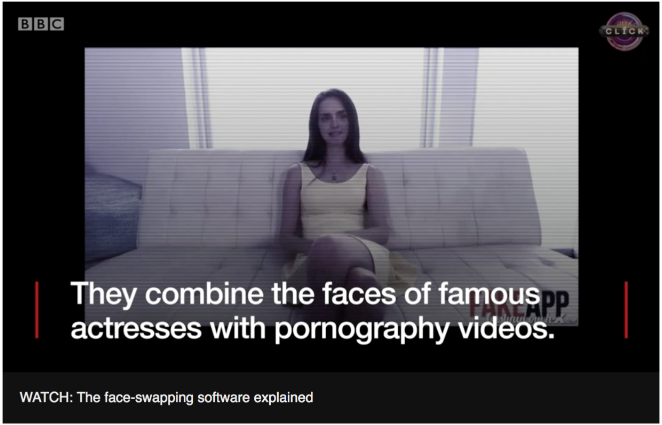
Class imbalanced datasets are common across different domains including health, security, banking and others. A typical supervised learning algorithm tends to be biased towards the majority class when dealing with imbalanced datasets. The learning task becomes more challenging when there is also an overlap of instances from different classes. In this paper, we propose an undersampling framework for handling class imbalance in binary datasets by removing potential overlapped data points. Our methods are designed to identify and eliminate majority class instances from the overlapping region. Accurate identification and elimination of these instances maximise the visibility of the minority class instances and at the same time minimises excessive elimination of data, which reduces information loss. Four methods based on neighbourhood searching with different criteria to identify potential overlapped instances are proposed in this paper. Extensive experiments using simulated and real-world datasets were carried out. Results show comparable performance with state-of-the-art methods across different common metrics with exceptional and statistically significant improvements in sensitivity.

# AI Performance

- Humans were outperformed by algorithms in many CV tasks
- On a lower quality images, humans and algorithms performance is similar [14]

[14] S. Dodge and L. Karam, "A Study and Comparison of Human and Deep Learning Recognition Performance under Visual Distortions," 2017 26th International Conference on Computer Communication and Networks (ICCCN), 2017, pp. 1-7, doi: 10.1109/ICCCN.2017.8038465.

- "Deepfake videos could 'spark' violent social unrest"[15]




[15] source https://www.bbc.co.uk/news/technology-48621452

# Data Authenticity

- It is now possible not only to produce convincing forged video but also to fully synthesize video content
- So many ways to edit/ forge a video or image
- No training data available!

Digital Investigation 29 (2019) 67–81

Contents lists available at ScienceDirect

Digital Investigation

journal homepage: www.elsevier.com/locate/diin

ELSEVIER

## A review of digital video tampering: From simple editing to full synthesis

Pamela Johnston*, Eyad Elyan

*Robert Gordon University, Garthdee House, Garthdee Road, Aberdeen, AB10 7QB, Scotland, UK*

A R T I C L E I N F O

*Article history:*
Received 16 November 2018
Received in revised form
24 January 2019
Accepted 17 March 2019
Available online 22 March 2019

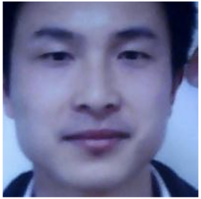
*Keywords:*
Video tampering
Video synthesis
Deep learning
Video forgery

A B S T R A C T

Video tampering methods have witnessed considerable progress in recent years. This is partly due to the rapid development of advanced deep learning methods, and also due to the large volume of video footage that is now in the public domain. Historically, convincing video tampering has been too labour intensive to achieve on a large scale. However, recent developments in deep learning-based methods have made it possible not only to produce convincing forged video but also to fully synthesize video content. Such advancements provide new means to improve visual content itself, but at the same time, they raise new challenges for state-of-the-art tampering detection methods. Video tampering detection has been an active field of research for some time, with periodic reviews of the subject. However, little attention has been paid to video tampering techniques themselves. This paper provides an objective and in-depth examination of current techniques related to digital video manipulation. We thoroughly examine their development, and show how current evaluation techniques provide opportunities for the advancement of video tampering detection. A critical and extensive review of photo-realistic video synthesis is provided with emphasis on deep learning-based methods. Existing tampered video datasets are also qualitatively reviewed and critically discussed. Finally, conclusions are drawn upon an exhaustive and thorough review of tampering methods with discussions of future research directions aimed at improving detection methods.

# Data Authenticity

- Face presentation attacks



Contents lists available at ScienceDirect

Information Fusion

journal homepage: www.elsevier.com/locate/inffus

A review of state-of-the-art in Face Presentation Attack Detection: From early development to advanced deep learning and multi-modal fusion methods

Faseela Abdullakutty *, Eyad Elyan, Pamela Johnston

*School of Computing, Robert Gordon University, Aberdeen, United Kingdom*

ARTICLE INFO

*Keywords:*
Face Recognition
Presentation attacks
Deep learning
Generalisation

ABSTRACT

Face Recognition is considered one of the most common biometric solutions these days and is widely used across a range of devices for various security purposes. The performance of FR systems has improved by orders of magnitude over the past decade. This is mainly due to the latest developments in computer vision and deep convolutional neural networks, and the availability of large training datasets. At the same time, these systems have been subject to various types of attacks. Presentation attacks are common, simple, and easy to implement. These simply involve presenting a video, photo, or mask to the camera or digital sensor and have proven capable of fooling FR systems and providing access to unauthorised users. Presentation attack detection is increasingly attracting more attention in the research community. A wide range of methods has already been developed to address this challenge. Deep learning-based methods in particular have shown very promising results. However, existing literature suggests that even with state-of-the-art methods, performance drops significantly in cross-dataset evaluation. We present a thorough, comprehensive, and technical review of existing literature on this timely and challenging problem. We first introduce and discuss the presentation attack problem and cover related and recent work in this area. In-depth technical details of existing presentation attack detection methods are then presented and critically discussed and evaluated, followed by a comprehensive discussion and evaluation of existing public datasets and commonly used evaluation metrics. Our review shows clearly that despite the recent and significant advances in this area of research, detecting unseen attacks is still considered a key problem. Machine learning methods tend to perform well, but only when test data comes from the same distribution as the training data (i.e. same dataset). New research directions are discussed in detail, including ways to improve the generalisation of machine learning methods, and move towards creating more stable presentation attack detection techniques that generalise across a wide range of unseen samples.

## Plan

1. Background
   - Computer Vision
   - Image Representation
2. Breakthrough in CV
3. Common CV Tasks
   - Basic Tasks
   - Advanced Tasks
4. Real World Projects
   - Engineering Diagrams
   - Online User's Authentication
   - Remote Inspection
   - Mechanical Engineering Diagrams
5. Challenges
6. Conclusion

## Conclusion and Way Forward

- Understanding your data, and finding and articulating a problem is the most important and crucial step for building intelligent machine vision solution

# Conclusion and Way Forward

- Understanding your data, and finding and articulating a problem is the most important and crucial step for building intelligent machine vision solution

- The key challenging problem -I think- is to understand the context

# Thank You

**@ElyanEyad**

https://www3.rgu.ac.uk/dmstaff/elyan-eyad

https://github.com/heyad/Teaching/tree/master/Python