Estimating poverty from aerial images: First results

Joep Burger
Statistics Netherlands

j.burger@cbs.nl

April 6, 2022, ONS-UNECE Machine Learning Group - Monthly Forum

Introduction

– UN Sustainable Development Goals
– Estimate poverty at local level
  - surveys
  - registers
  - earth observation
– CBS: complete enumeration of income-related poverty
  - skip annotation
  - experiments
  - extract features for models

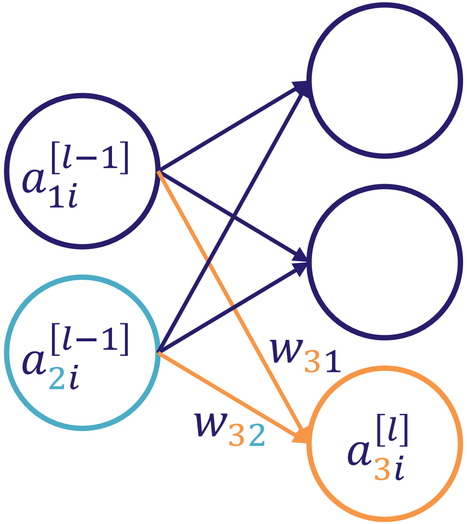
DeepGeoStat

EU grant on integration of geospatial information and statistics (2021–2022)

- WP1: user-friendly tool
- **WP2: poverty**
- WP3: land use
- WP4: land quality and biodiversity
- WP5: solar panels
- WP6: coordination and dissemination

Artificial neural network

$$\mathbf{A}^{[0]} = \mathbf{X}$$
$$\mathbf{A}^{[l]} = g\left(\mathbf{W}^{[l]}\mathbf{A}^{[l-1]} + \mathbf{b}^{[l]}\mathbf{1}\right)$$
$$\mathbf{A}^{[L]} = \mathbf{Y}$$

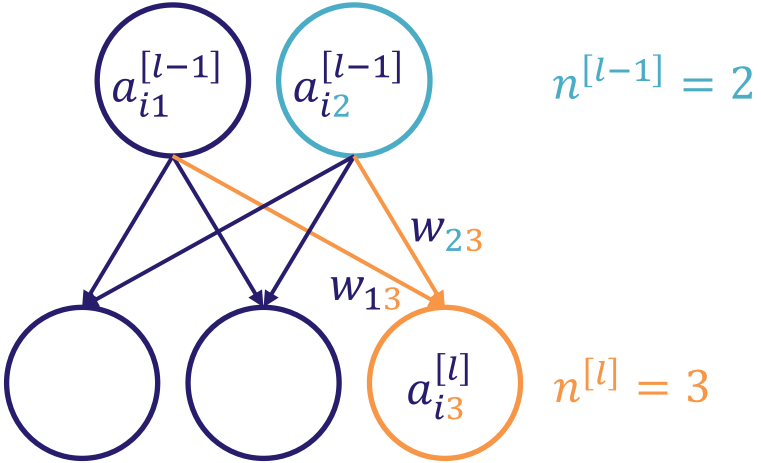


$n^{[l-1]} = 2$ $n^{[l]} = 3$

$$\begin{bmatrix} a_{11} & \cdots & a_{1m} \\ a_{21} & \cdots & a_{2m} \\ a_{31} & \cdots & a_{3m} \end{bmatrix}^{[l]} = g\left(\begin{bmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \\ w_{31} & w_{32} \end{bmatrix}^{[l]} \begin{bmatrix} a_{11} & \cdots & a_{1m} \\ a_{21} & \cdots & a_{2m} \end{bmatrix}^{[l-1]} + \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}^{[l]} \begin{bmatrix} 1 & \cdots & 1 \end{bmatrix} \right)$$

Alternative notation and visualization

$$\mathbf{A}^{[0]} = \mathbf{X}$$
$$\mathbf{A}^{[l]} = g\left(\mathbf{A}^{[l-1]}\mathbf{W}^{[l]} + \mathbf{1}\mathbf{b}^{[l]}\right)$$
$$\mathbf{A}^{[L]} = \mathbf{Y}$$



$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & a_{m3} \end{bmatrix}^{[l]} = g\left(\begin{bmatrix} a_{11} & a_{12} \\ \vdots & \vdots \\ a_{m1} & a_{m2} \end{bmatrix}^{[l-1]} \begin{bmatrix} w_{11} & w_{12} & w_{13} \\ w_{21} & w_{22} & w_{23} \end{bmatrix}^{[l]} + \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \begin{bmatrix} b_1 & b_2 & b_3 \end{bmatrix}^{[l]}\right)$$

Learning weights

- Minimize cost function

$$J(w,b) = \frac{1}{m}\sum_{i=1}^{m} L\left(\hat{y}^{(i)}, y^{(i)}\right) + \lambda R(w)$$

- Initialize weights
- Update iteratively using partial derivatives and learning rate η

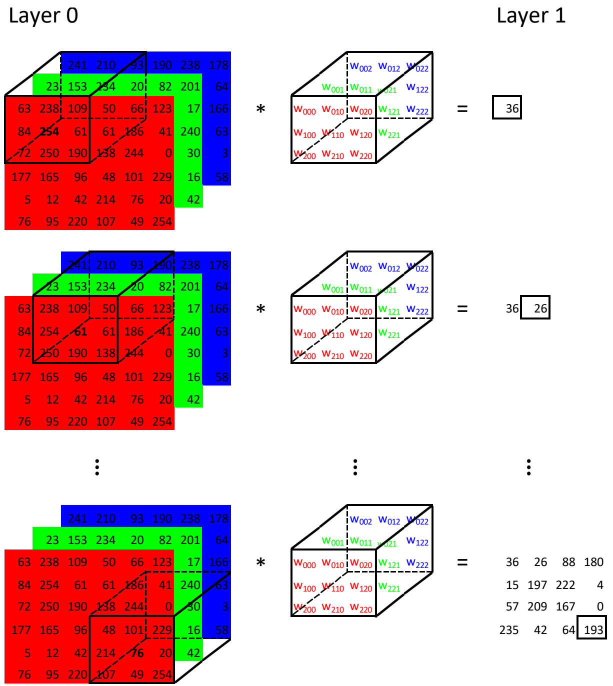
$$w_j \coloneqq w_j - \eta \frac{\partial J(w,b)}{\partial w_j}$$

Convolutional neural networks

- Convolutional layers
- Spatial context
- Reduce number of parameters

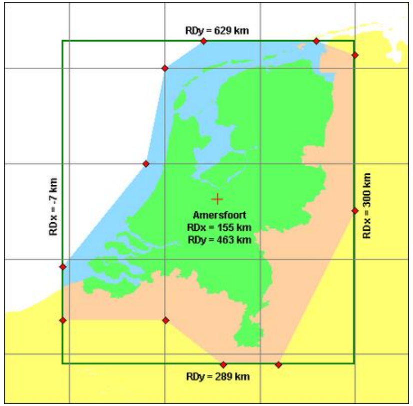
  FC: $(400 \times 400 \times 3 = 480{,}000) \times n^{[1]}$

  CONV: $(3 \times 3 \times 3 = 27) \times n_c^{[1]}$



Data: grid

– Coordinate Reference System Amersfoort
– R = 100: 3.5 mln 1-ha squares
– R = 500: 144 thousand 25-ha squares

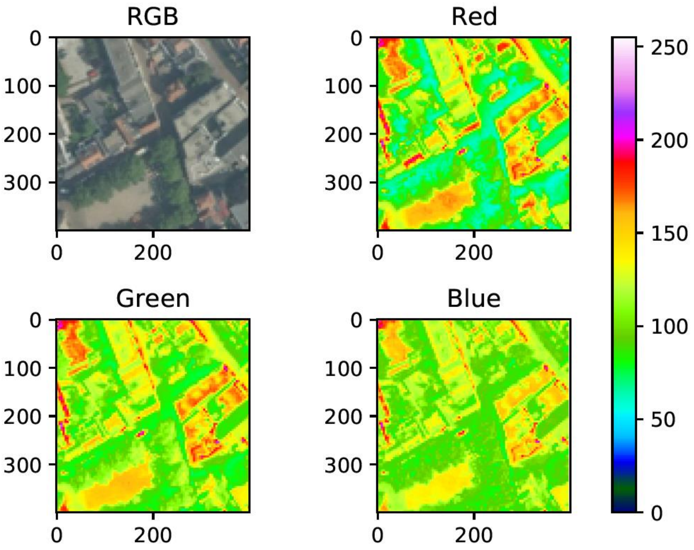


wikipedia

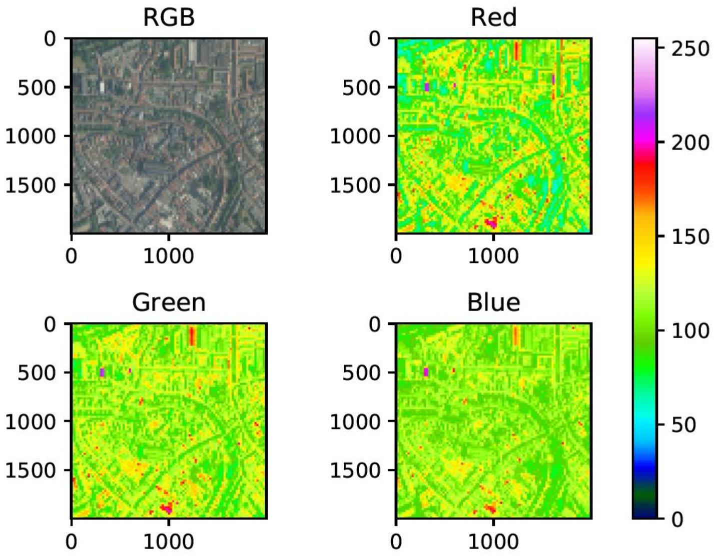
Data: images

– Aerial, 0.25 m per pixel, RGB, 8-bit color palette, summer, pdok.nl

R = 100



R = 500



Data: labels

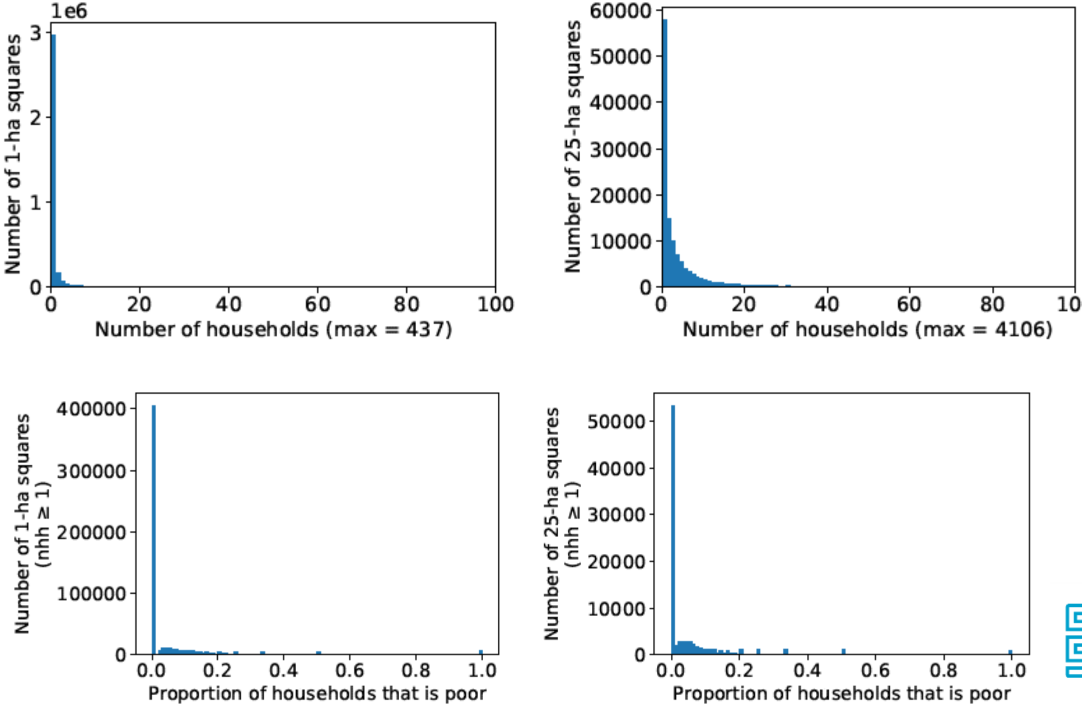
- Complete enumeration from registers
- About 8 mln private households (2017)
- Geotagged standardized disposable household income
- Poor: below low-income threshold (social security rate single person, i.e. 12,480 euro per year in 2017)
- Mapped to square

Data: labels

– Indicator for inhabitation

– Indicator for at least 1 poor household given inhabited

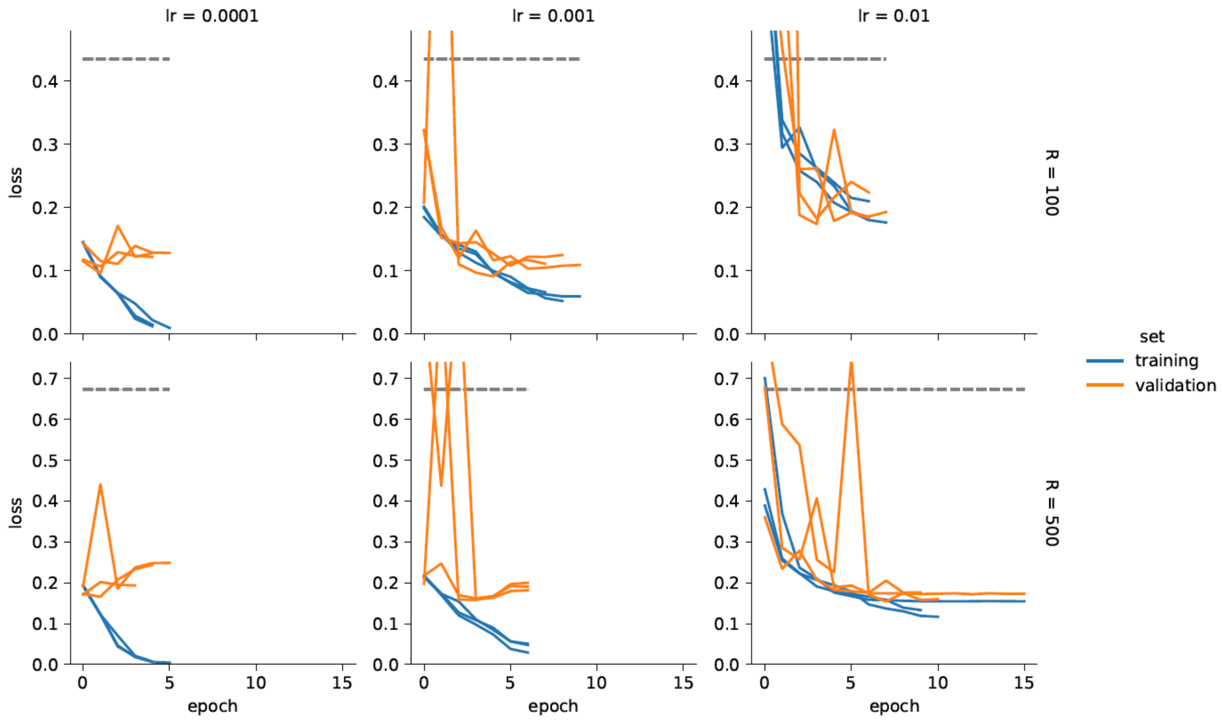
R = 100 R = 500



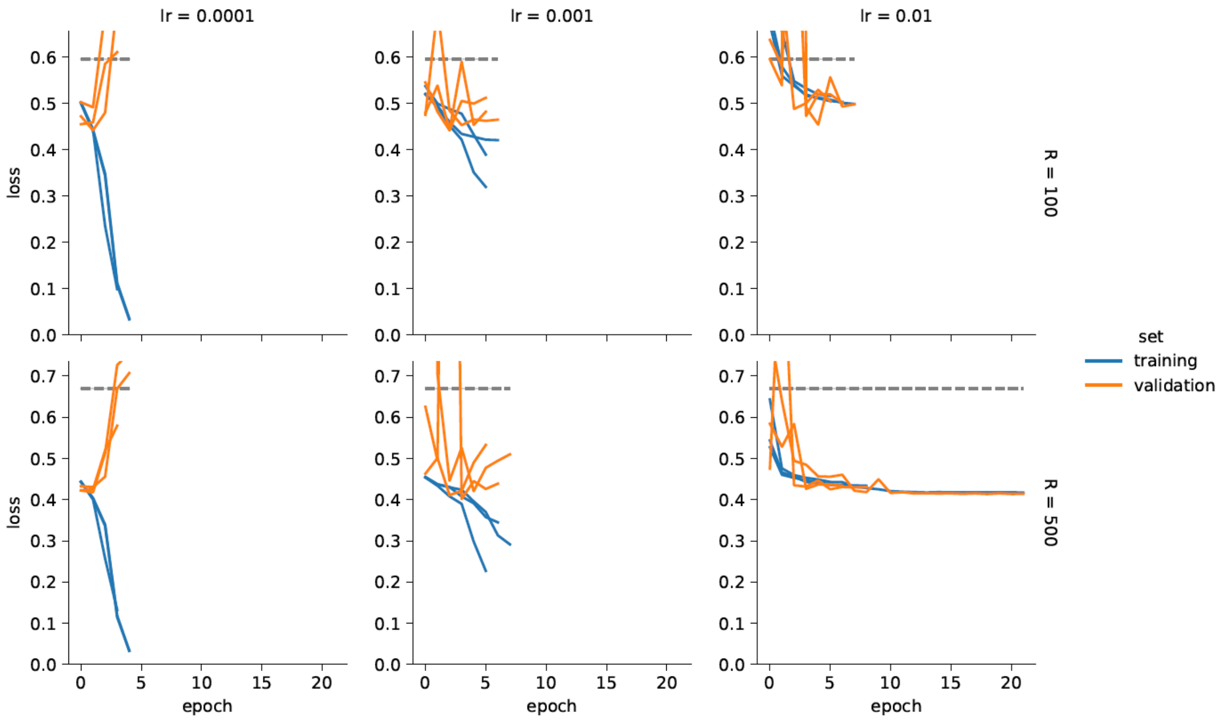
Implementation

- Python library Keras
- InceptionV3 architecture
- Pretrained on 1.2 mln ImageNet images, 1000 categories
- Final layer: 1000 nodes, softmax → 1 node, sigmoid
- All layers trainable
- $\frac{1}{10}$ test set, simple random sample of max 10^5, $\frac{1}{10}$ validation set
- Images resized to default 299 × 299 × 3 using bilinear interpolation
- Adam optimizer
- Batch size 128
- Class weights inversely proportional to prevalence in training set
- 32 GB Nvidia Tesla V100 GPU in DGX-1 server

Results: inhabitation



Results: poverty given inhabited



Summary

- Models learn to classify images better than guessing
- Inhabitation easier than poverty
- 25-ha squares easier than 1-ha squares
- Without regularization, higher learning rate prevented overfitting

Discussion

- Reproducibility
- Quality metrics (ACC, AUC, GM, PSS, MCC, MRK, F1, ...)
- When good enough?
- Explainable AI

Future

- Regularization
  - penalize loss function
  - apply dropout to FC layers
  - data augmentation
- Near infrared (vegetation)
- Incorporate label quality (e.g. number of households)
- Ordinal target variable (e.g. none, some, many)
- Effects of sample size, sampling design and image resolution
- Auxiliary labels (e.g. urbanicity)
- Activation maps
- Feature extraction

References

Burger, J., N. de Wolf and J. van den Brakel (2022). Deep Geo Stat, work package 2: estimating poverty from earth observation, deliverable 2.1: intermediate report. Eurostat.

Burger, J. and Q. Meertens(2020). The algorithm versus the chimps: On the minima of classifier performance metrics. *In* Proceedings of the BNAIC/BeneLearn.

Jean, N., M. Burke, M. Xie, W.M. Davis, D.B. Lobell and S. Ermon (2016). Combining satellite imagery and machine learning to predict poverty. *Science* 353(6301), 790–794.

Perez, A., C. Yeh, G. Azzari, M. Burke, D. Lobell and S. Ermon (2017). Poverty prediction with public Landsat7 satellite imagery and machine learning. *arXiv* 1711.03654.

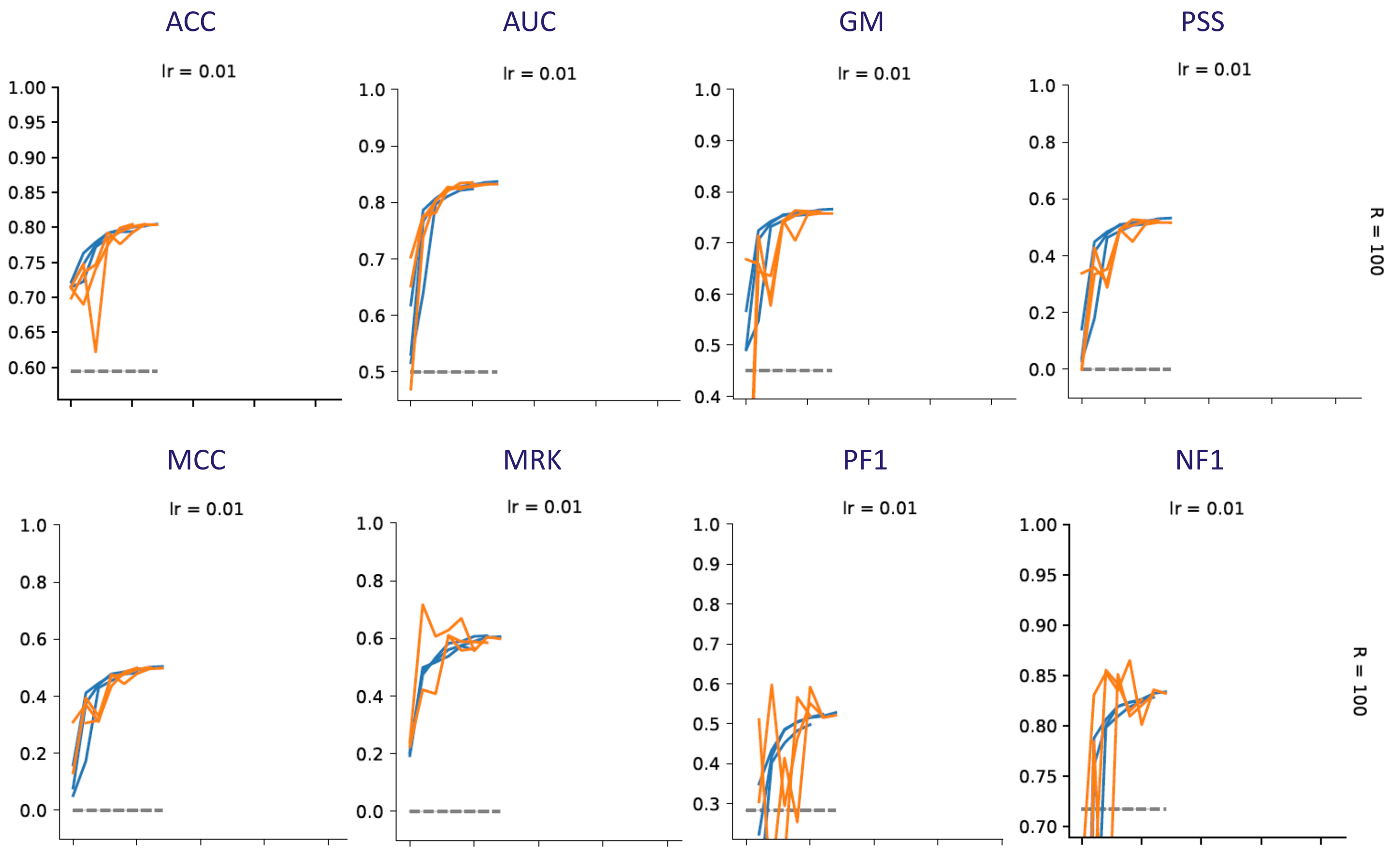
Rosebrock, A. (2018). *Deep Learning for Computer Vision* (2nd ed.). PyImageSearch.

Szegedy, C., V. Vanhoucke, S. Ioffe, J. Shlens and Z. Wojna (2016). Rethinking the inception architecture for computer vision. *In* Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

Yeh, C., A. Perez, A. Driscoll, G. Azzari, Z. Tang, D. Lobell, S. Ermon and M. Burke (2020). Using publicly available satellite imagery and deep learning to understand economic well-being in africa. *Nature Communications* 11(2583).

Thank you

pdok.nl


ACC
AUC
GM
PSS
MCC
MRK
PF1
NF1
lr = 0.01
R = 100