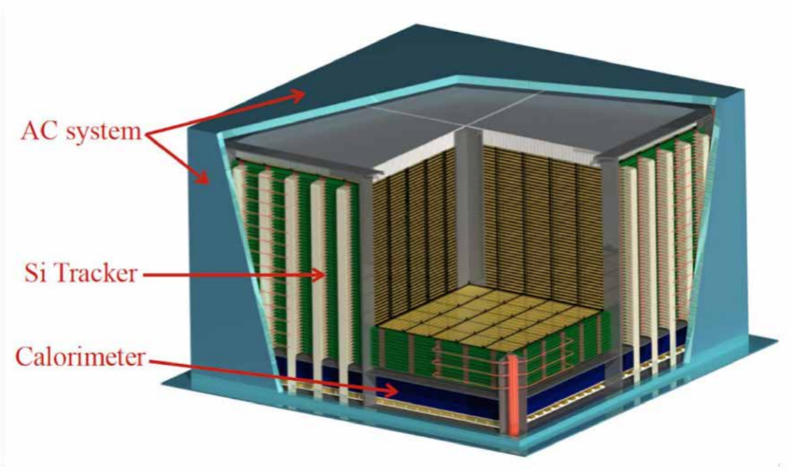
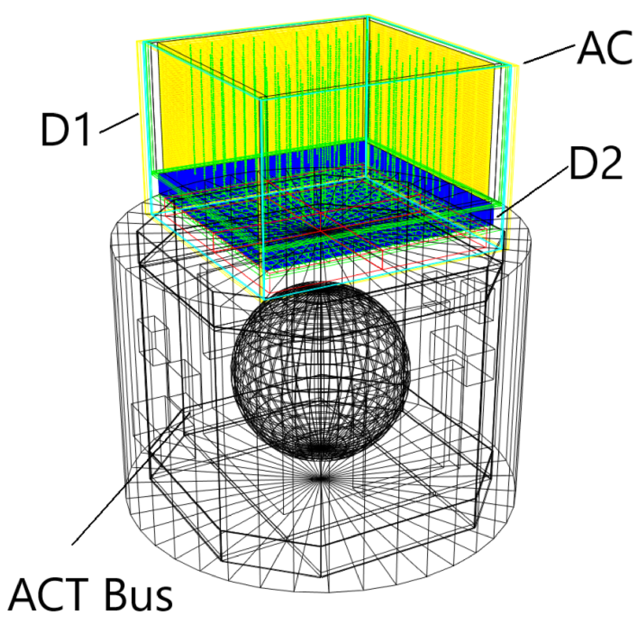
CNNCat: Categorizing high-energy photons in a Compton/Pair Telescope with convolutional neural networks[1]

Jan Peter Lommler, Uwe Oberlack (Johannes Gutenberg Universität Mainz & Exzellenzcluster Prisma++)

A Compton/Pair telescope, designed to provide spectrally resolved images of cosmic photons from sub-MeV to GeV energies, records a wealth of data in a combination of tracking detector and calorimeter. Onboard event classification can be required to decide on which data to downlink with priority, given limited data-transfer bandwidth. Event classification is also the first and one of the most crucial steps in reconstructing data. Its outcome determines the further handling of the event, i.e., the type of reconstruction (Compton, pair) or, possibly, the decision to discard it. Errors at this stage result in misreconstruction and loss of source information. We present a classification algorithm driven by a Convolutional Neural Network. It provides classification of the type of electromagnetic interaction, based solely on low-level detector data. We introduce the task, describe the architecture, and present the performance of this method in the context of the proposed (e-)ASTROGAM and similar telescopes.

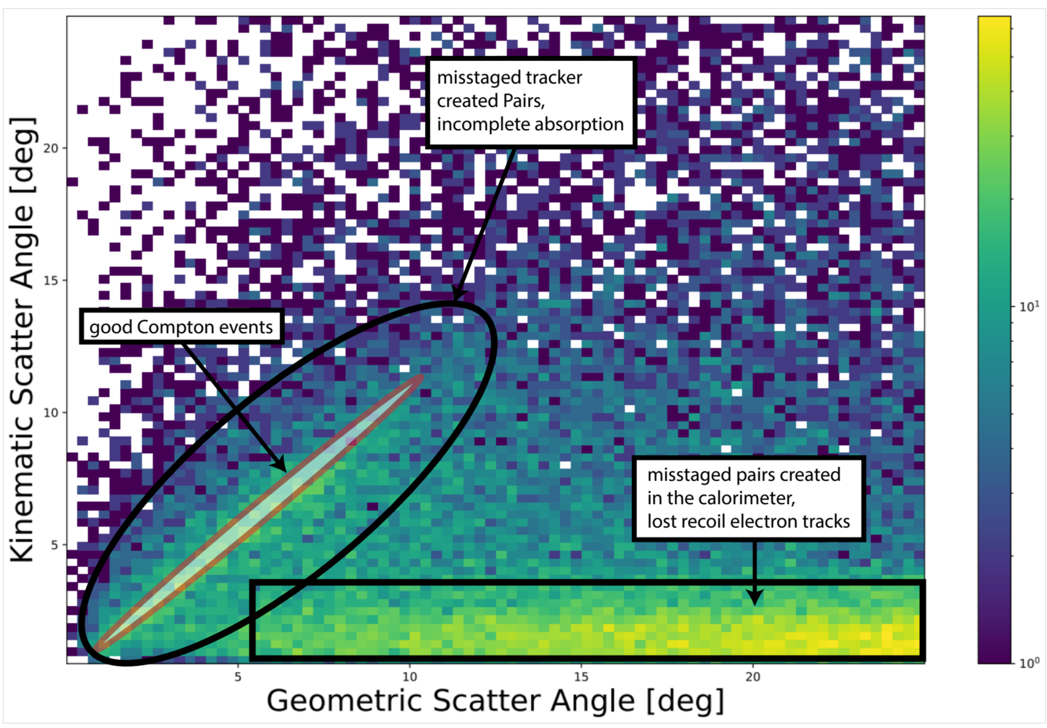
Compton/Pair Telescopes





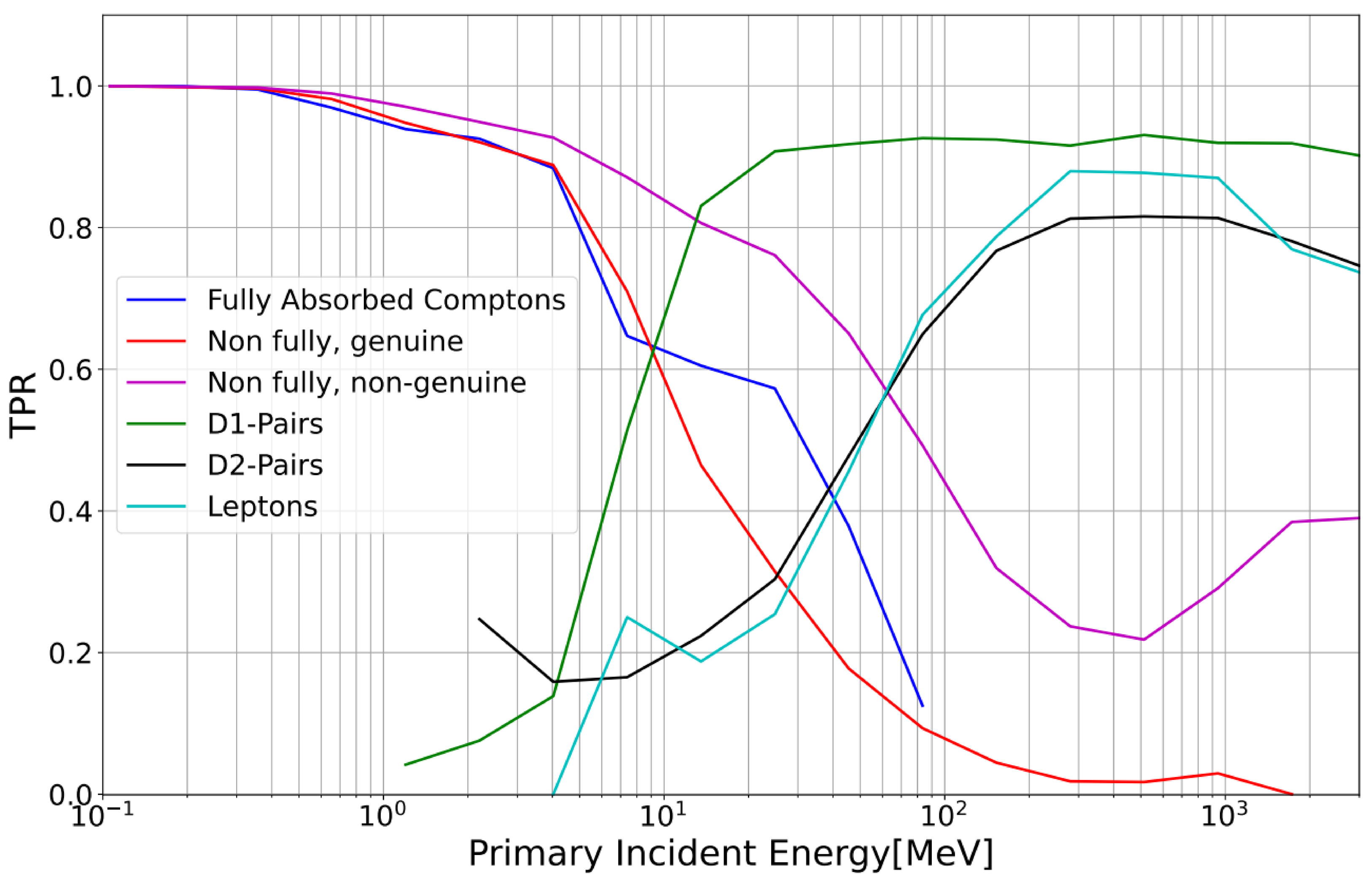
- Compton/Pair telescopes are designed to measure over the range of 100 keV to several GeV
- Typical designs consist of stacked Silicon detectors (either DSSDs or pixel detectors) as tracker (D1) and a solid state calorimeter as absorber (CsI, NaI, $CeBr_3$) (D2).
- Up to 10 MeV, the telescopes detects photons mainly via Compton scatter.
- Pair creation is the dominant detection mode for higher incident energies.
- In order to protect the detector from charged primary and secondary cosmic rays, the setup is surrounded by an anti-coincidence veto (AC) made of plastic scintillator.
- e-ASTROGAM[2] was envisioned with 56 layers of DSSDs as D1 and small CsI finger crystals as D2 covering an area of roughly 1 m^2.

Why is classification needed?



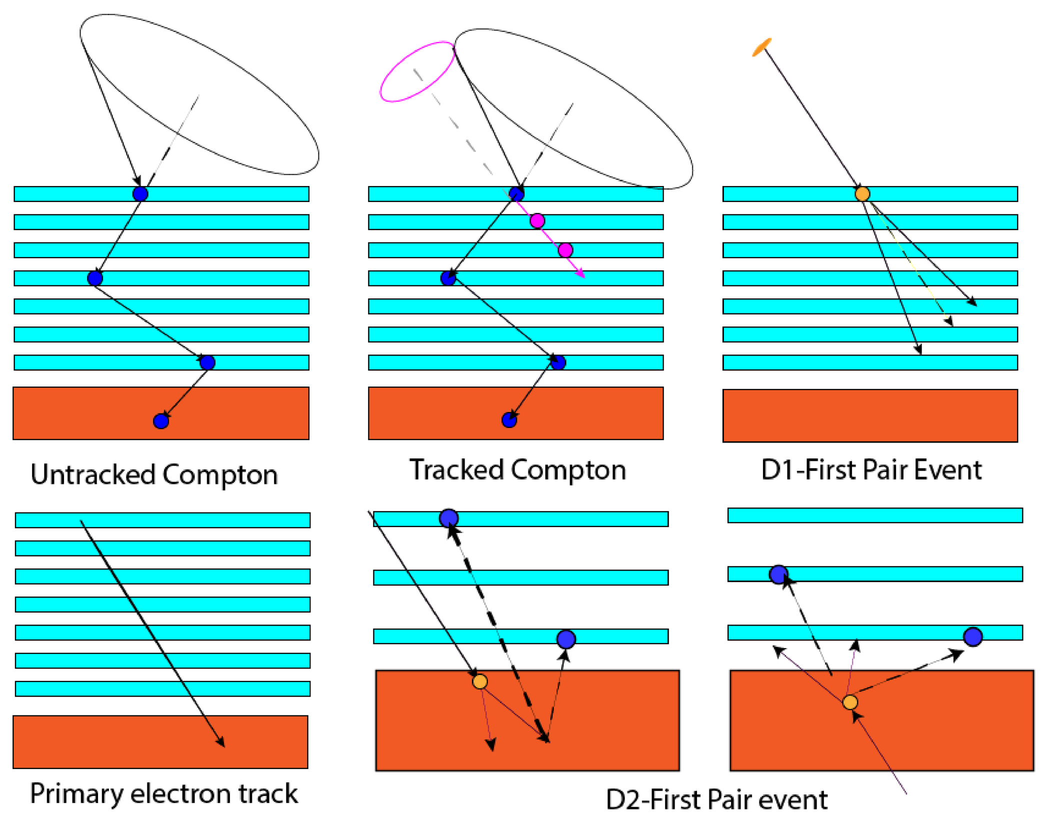
- Data reconstructed without prior classification leads to mistagged events wrongfully reconstructed as Compton events which leave imprints in, e.g., the reconstructed vs geometric scatter angle space (8.5 MeV on-axis MEGAlib[3] simulation)
- Better tagging of D1-pairs and D2-pairs prior to reconstruction would remove parts of the halo and the band structure at low kinematic scatter angles.

Class Coverage



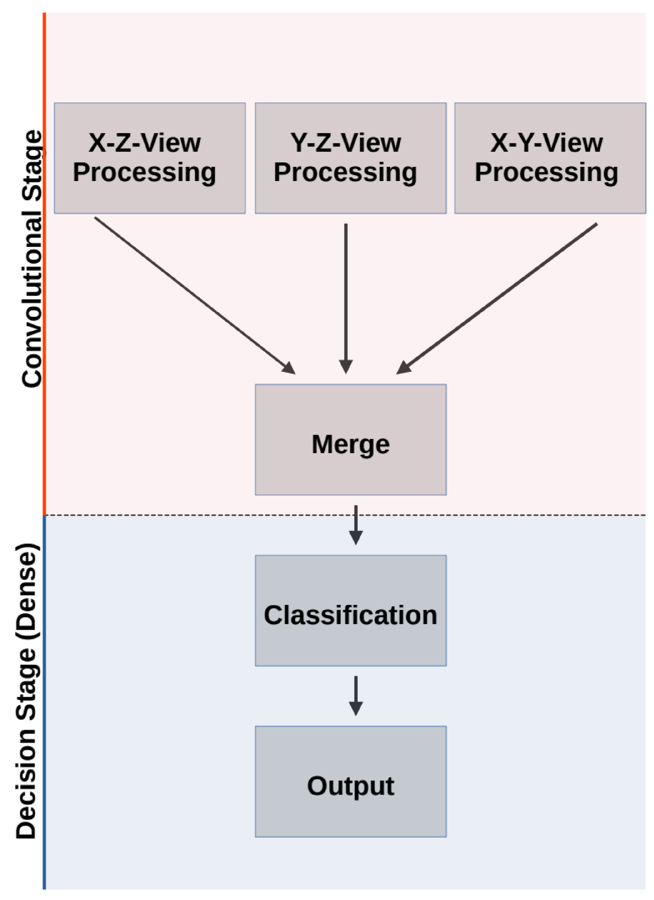
- The test set is composed out of 2.5 M unseen samples in the range of 100 keV to 3.5 GeV and 0-180 degrees incident angle equally binned in energy and incident angle.
- The final model recognizes the learned classes very well.
- Compton events are tagged reliably up to 10 MeV.
- D1-pair events are tagged at or above 90% 20 MeV.
- D2-pairs and charged primaries are tagged reliably above 60 MeV.

Events in Compton/Pair Telescopes



- Compton scatters can occur with and without visible recoil electron track (tracked and untracked Compton event).
- For untracked Compton events, the origin of the photon can be restricted to a circle on the sky.
- For tracked Compton events, the recoil electron track allows to restrict the circle to an arc.
- Electron-Positron pairs created in the tracker (D1-first pairs) can be backprojected to the origin of the photon.
- Pairs created in the calorimeter can create hit patterns resembling Compton and D1-first pairs, while not containing any relevant information about the source.
- Charged particles leaking in can create track-like patterns that could resemble legit events. In contrast to D2-pairs, they are a genuine background component.

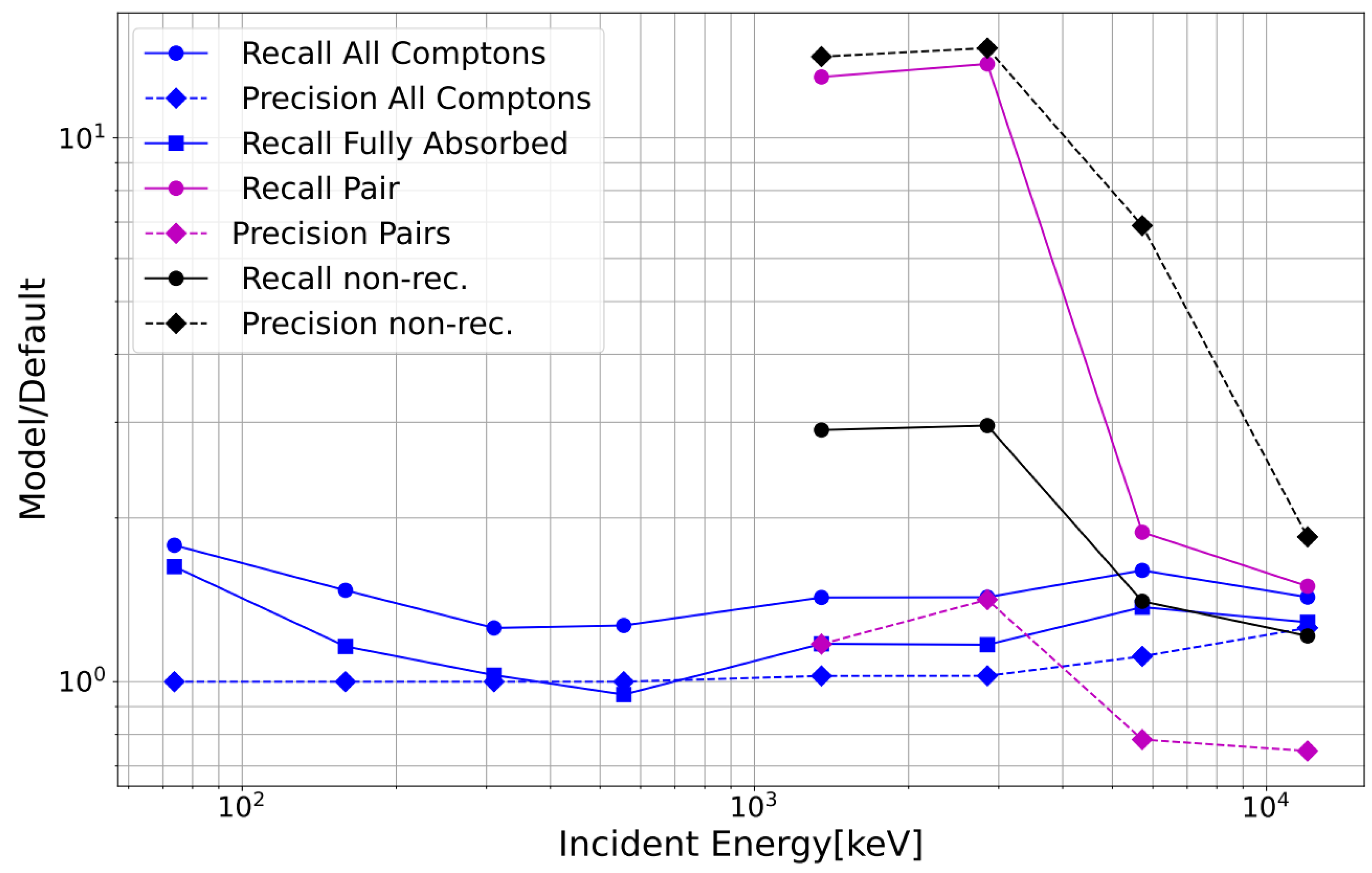
Model Architecture



- A full 3D image of an event is in the order of 1 GB for e-ASTROGAM, 3D convolutions would demand a large amount of VRAM.
- The issue can be circumvented this by using 2D projections along X-Y, X-Z and Y-Z with around 2 MB size each.
- Convolutional layers extract the features, the dense layer(s) classifies the event.
- The very lightweight architecture is driven by the SqueezeNet microarchitecture[4] , depthwise separable convolutions[5] and spatial pyramid pooling[6].
- Only 64k trainable weights (typically > 1 million)
- Edge computing hardware like an NVIDIA Jetson AGX Orin deliver 1.3 TFLOPS/Watt (FP16) or 2.6 TOPS/Watt (INT8), which allows up to 1.4 kHz@FP16 or 2.9 kHz@INT8 events per Watt

This model can run on-board on dedicated hardware!

Improvement



- Tagging accuracy is improved compared to the default MEGAlib classification-by-reconstruction
- Separation of low energetic D1-pair events from Compton events is improved significantly
- Compton tagging is improved over the whole range
- D2-first pairs are also removed more efficiently

Author affiliations:
Institut für Physik,
Johannes Gutenberg-
Universität Mainz

With funding from the:
Federal Ministry of Research, Technology and Space

DLR

This work is supported by the German Aerospace Center (DLR) under grant number 50OO2218.

References:

[1] Exp Astron58, 18 (2024) , https://doi.org/10.1007/s10686-024-09965-5
[2] Exp. Astron. 44, 25 (2017), The e-ASTROGAM mission, https://doi.org/10.1007.s10686-017-9533-6
[3 NewAR 50, 629 (2006), MEGAlib, https://doi.org/10.1016/j.newar.2006.06.049]
[4] eprint arXiv:1602.07360: SqueezeNet
[5] eprint arXiv:1610.02357, Xception: Deep Learning with Depthwise Separable Convolutions
[6] European Conference on Computer Vision 2014, Spatial Pyramid Pooling, https://doi.org/10.1007/978-3-319-10578-9_23