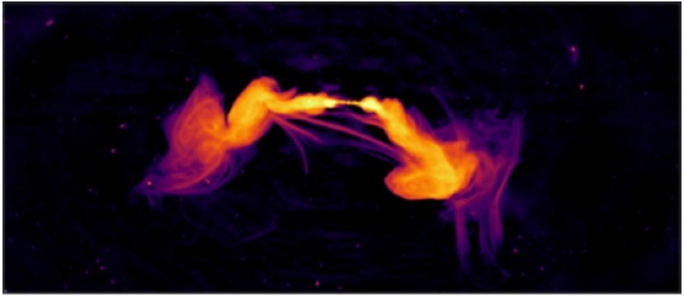
technische universität
dortmund

e5 experimentelle physik 5
astroteilchenphysik



Analysis of MOJAVE data with Neural Networks

Arne Poggenpohl

December 16, 2022

Astroparticle Physics / WG Elsässer
Department of Physics, TU Dortmund University

Cosmic messengers

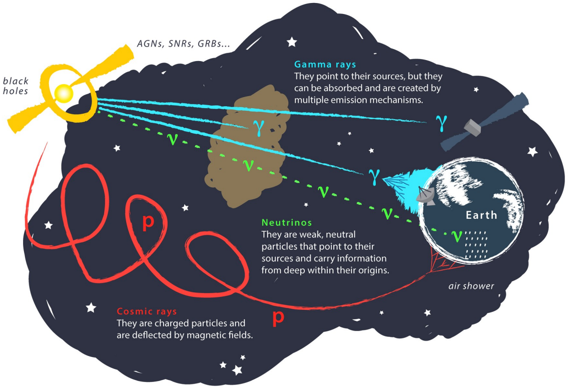


Figure: Different types of messengers in astronomy. [1]

Radio astronomy

- Atmosphere is transparent for radio radiation
- Rayleigh criterion:

$$\theta \approx 1{,}22 \cdot \frac{\lambda}{D}$$

$$\text{i.e.} \quad \theta \approx 1{,}22 \cdot \frac{0{,}21\,\mathrm{m}}{25\,\mathrm{m}} = 0{,}0084 \approx 0{,}48°$$

- Single telescope size constrained

→ Combining array of telescopes

  → Radio interferometry

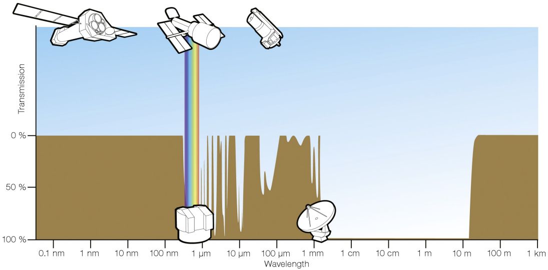
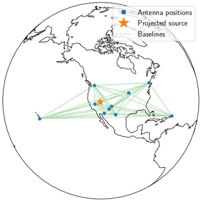
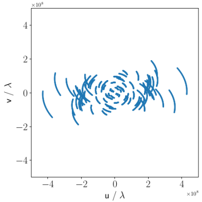


Figure: Transmission of radiation in the atmosphere. [2]

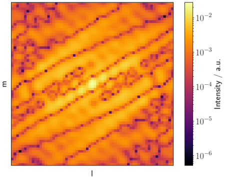
Radio interferometry



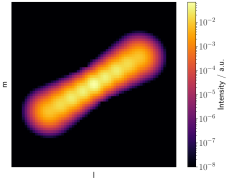
(a) The Very Long Baseline Array. [3]



(b) Limited (u, v) coverage. [3]



(c) Dirty image after FFT. [3]



(d) Source in spacial space. [3]

MOJAVE data

- Monitoring Of Jets in Active galactic nuclei with VLBA Experiments
- Measurement series over several years
- Very Long Baseline Array
  - 10 telescopes
  - 25 m diameter
  - 8611 km maximum baseline
  - 0,17 – 22 mas angular resolution
- Similar data quality throughout the data set

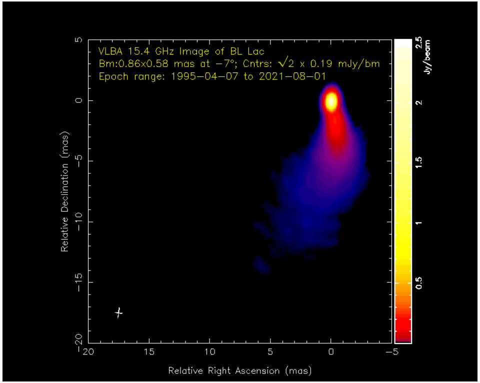


Figure: Source in the MOJAVE dataset. [4]

Problems

- Jets consist of several components (bursts)
- Track components over time to determine kinematics

→ Physical properties of the host galaxy

Problems

- Jets consist of several components (bursts)
- Track components over time to determine kinematics

→ Physical properties of the host galaxy

Solutions

- Previously: `PyBDSF` or `DIFMAP`
  - Analysis performed manually

  → Takes long for big data

  → Difficult to reproduce

Problems

- Jets consist of several components (bursts)
- Track components over time to determine kinematics

→ Physical properties of the host galaxy

Solutions

- Previously: `PyBDSF` or `DIFMAP`
  - Analysis performed manually

  → Takes long for big data

  → Difficult to reproduce
- This approach:
  - Creates simulations
  - Uses neural network to track components

  → Building and training neural network requires expertise

  → Fast evaluation, reproducible results

Neural network for object detection

- Object detection is famous task for NN
- Convolutional layers to extract features

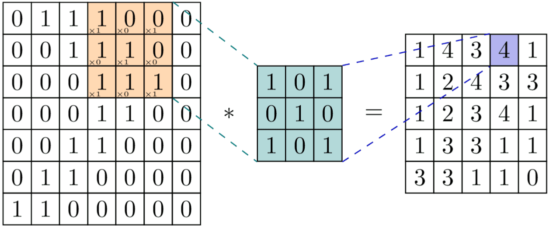


Figure: A 2D Convolution. [5]

Neural network for object detection

- Object detection is famous task for NN
- Convolutional layers to extract features

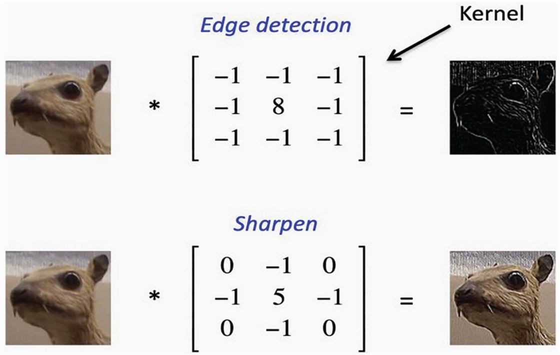


Figure: Application examples of kernels. [6]

Neural network for object detection

- Object detection is famous task for NN
- Convolutional layers to extract features
- Different approaches for object detection
  - Evaluate classifier at several locations and scales
  - Sliding window approach
  - R-CNN:
    1. Predict potential boxes
    2. Apply classifier
    3. Post-processing

  → Complex pipelines are slow and hard to optimize

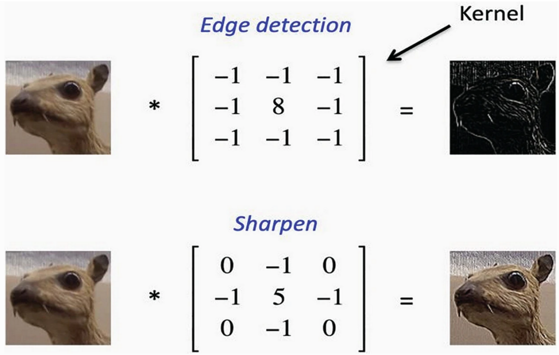


Figure: Application examples of kernels. [6]

Neural network for object detection

- Object detection is famous task for NN
- Convolutional layers to extract features
- Different approaches for object detection
  - Evaluate classifier at several locations and scales
  - Sliding window approach
  - R-CNN:
    1. Predict potential boxes
    2. Apply classifier
    3. Post-processing
  - → Complex pipelines are slow and hard to optimize
- → YOLO

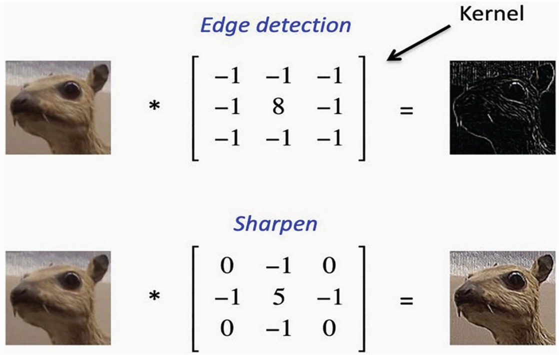


Figure: Application examples of kernels. [6]

You Only Look Once

- Predicts boxes, objectness and classes
- Every pixel in output is prediction
- Non-maximum suppression (NMS) to iteratively remove lower scoring boxes
- Layers
  - Convolution
  - Deconvolution
  - Activations: ReLU, LeakyReLU, SiLU
  - Batch normalization, MaxPooling

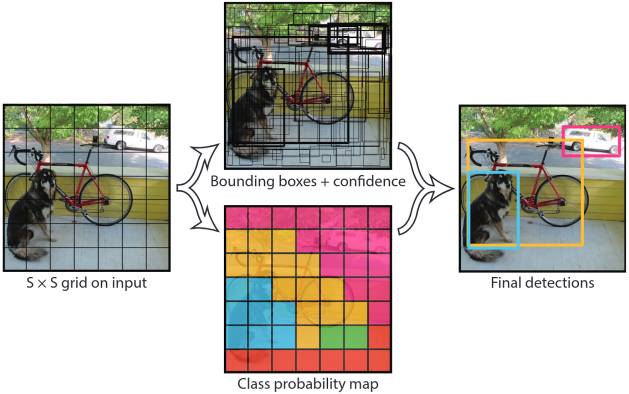


Figure: You Only Look Once: Unified, Real-Time Object Detection. (2015) [7]

YOLOv6 framework

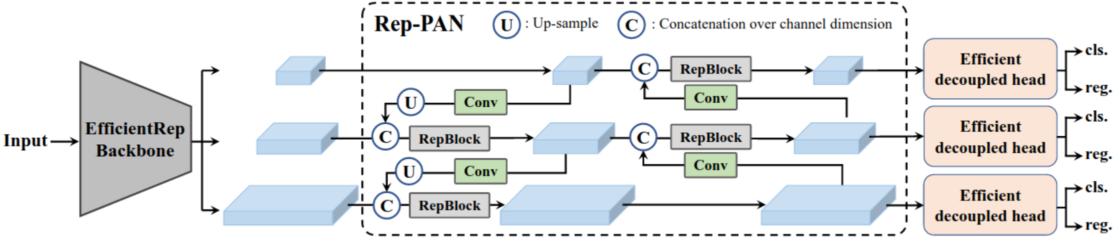


Figure: The YOLOv6 framework. [8]

Loss function

- Box regression: Complete Intersection over Union (`CIoU`)

$$L_{Box} = 1 - IoU + \frac{d^2}{c^2} + \alpha v$$

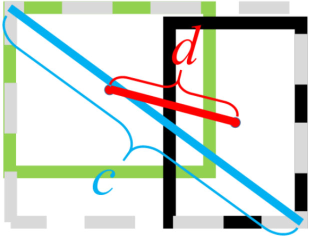


Figure: DIoU loss for bounding box regression. [9]

Loss function

- Box regression: Complete Intersection over Union (`CIoU`)

$$L_{Box} = 1 - IoU + \frac{d^2}{c^2} + \alpha v$$

- Objectness: Binary cross entropy with sigmoid (`BCEWithLogits`)

$$L_{Obj} = -w\left[y \cdot \log(\sigma(x)) + (1 - y) \cdot \log(1 - \sigma(x))\right]$$

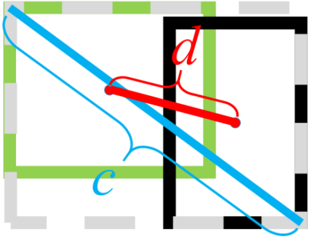


Figure: DIoU loss for bounding box regression. [9]

Loss function

- Box regression: Complete Intersection over Union (`CIoU`)

$$L_{Box} = 1 - IoU + \frac{d^2}{c^2} + \alpha v$$

- Objectness: Binary cross entropy with sigmoid (`BCEWithLogits`)

$$L_{Obj} = -w\left[y \cdot \log(\sigma(x)) + (1-y) \cdot \log(1-\sigma(x))\right]$$

- Rotation: Mean absolute error (`MAE` or `L1`)

$$L_{Rot} = |x - y|$$

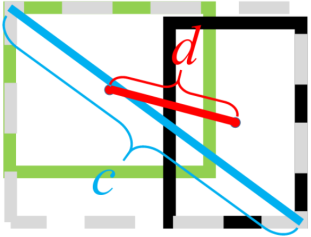


Figure: DIoU loss for bounding box regression. [9]

Loss function

- Box regression: Complete Intersection over Union (`CIoU`)

$$L_{Box} = 1 - IoU + \frac{d^2}{c^2} + \alpha v$$

- Objectness: Binary cross entropy with sigmoid (`BCEWithLogits`)

$$L_{Obj} = -w\left[y \cdot \log(\sigma(x)) + (1-y) \cdot \log(1-\sigma(x))\right]$$

- Rotation: Mean absolute error (`MAE` or `L1`)

$$L_{Rot} = |x - y|$$

- Total loss:

$$L = L_{Box} + L_{Obj} + L_{Rot}$$

- Training on **20000** simulated sources

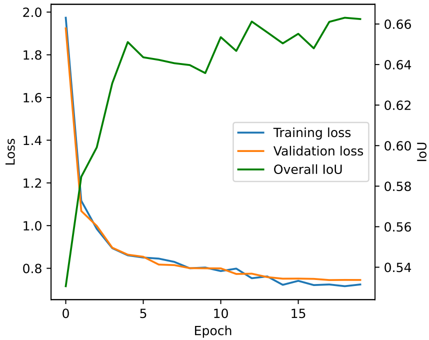


Figure: Loss and metric during training. [10]

Non-maximum supression

$$A: \text{List of predictions}, B: \text{Final list}$$
$$c: \text{IoU threshold}, d: \text{Objectness threshold}$$

1. Select box from A with highest objectness score (larger than d) and add it to B. (Initially B is empty).
2. Compare this predictions with all the predictions — calculate the IoU of this predictions all other predictions in B. If IoU is larger than c, remove that predictions from B.
3. This process is repeated until there are no more proposals left in A.

Simulation

- Gaussian distributions forming a jet
- Improvements:
  - Rotation of components
  - Larger eccentricity possible
  - Offset of core component
  - Randomly drop components
  - Advanced noise

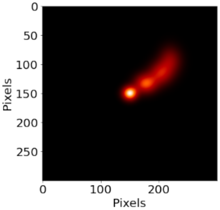


Figure: Simulated sources with `radiosim` in 2021. [11]

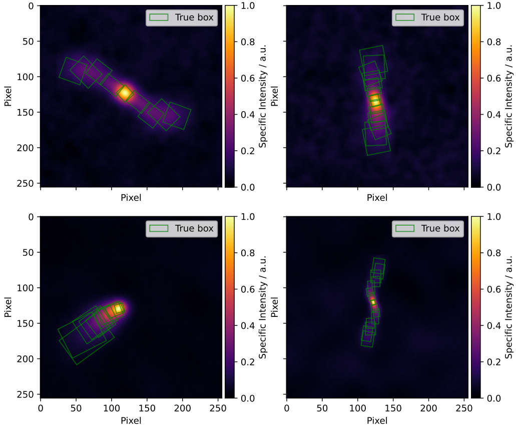


Figure: Simulated sources with `radiosim` for this work. [12]

Predictions for simulation data

- Output layers of size **64x64**, **32x32**, **16x16**
- Combine objectness by multiplication
- Boxes of largest output layer used for NMS
  - → **4096** pixels with predictions
  - → Thresholds in NMS determining

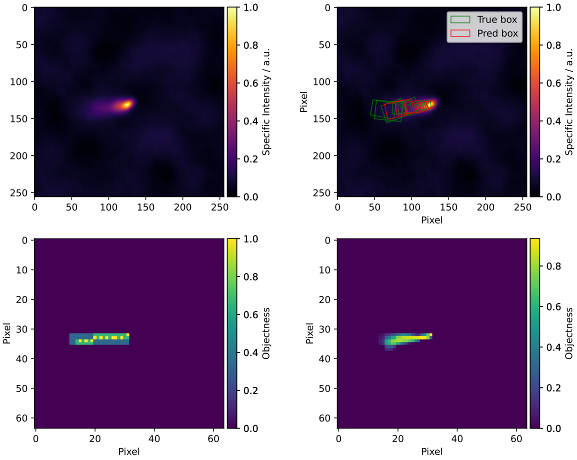


Figure: Predicted boxes and objectness.

Predictions for simulation data

- Output layers of size **64x64**, **32x32**, **16x16**
- Combine objectness by multiplication
- Boxes of largest output layer used for NMS
  - → **4096** pixels with predictions
  - → Thresholds in NMS determining

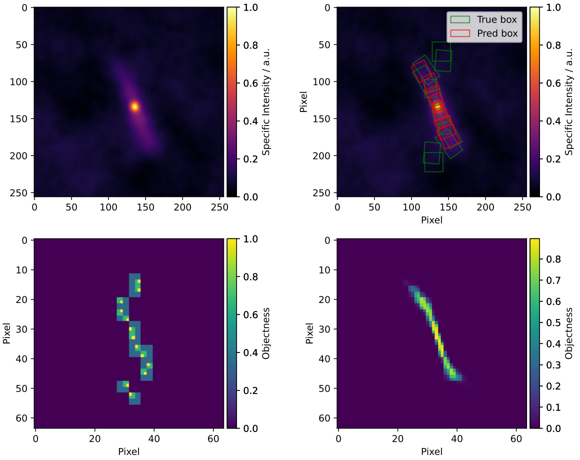


Figure: Predicted boxes and objectness.

Predictions for simulation data

- Output layers of size **64x64**, **32x32**, **16x16**
- Combine objectness by multiplication
- Boxes of largest output layer used for NMS
  - → **4096** pixels with predictions
  - → Thresholds in NMS determining

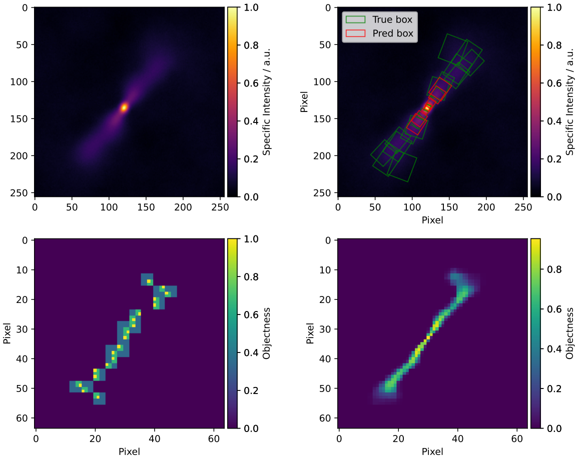


Figure: Predicted boxes and objectness.

Preprocessing of MOJAVE data

- Source is barely visible

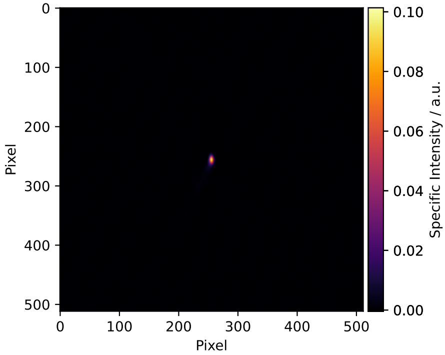


Figure: MOJAVE image of 1142+198 from 26.09.2016. [13]

Preprocessing of MOJAVE data

- Source is barely visible
- Zoom on source → Cropping to **128** pixel

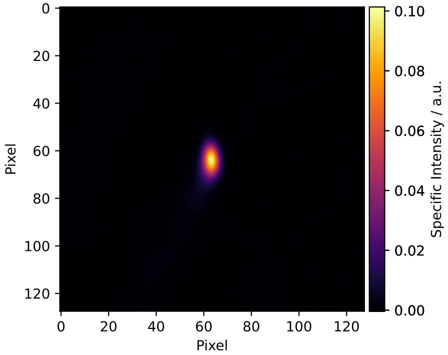


Figure: MOJAVE image of 1142+198 from 26.09.2016. [13]

Preprocessing of MOJAVE data

- Source is barely visible
- Zoom on source → Cropping to 128 pixel
- Scaling → $\log_{10}$

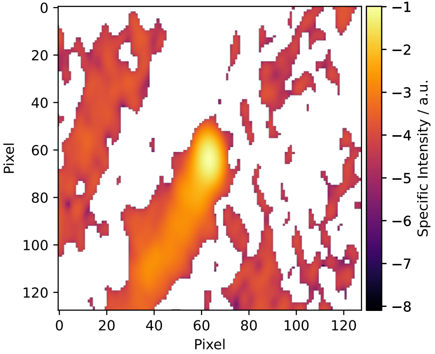


Figure: MOJAVE image of 1142+198 from 26.09.2016. [13]

Preprocessing of MOJAVE data

- Source is barely visible
- Zoom on source → Cropping to **128** pixel
- Scaling → $\log_{10}$
- Scaling → between **0** and **1**

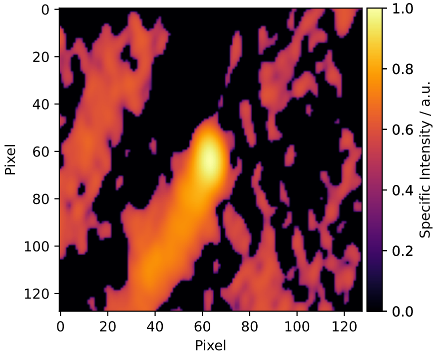


Figure: MOJAVE image of 1142+198 from 26.09.2016. [13]

Preprocessing of MOJAVE data

- Source is barely visible
- Zoom on source → Cropping to 128 pixel
- Scaling → $\log_{10}$
- Scaling → between 0 and 1
- Noise cut (before $\log_{10}$) → $x[x < |\min(x)|] = -1$

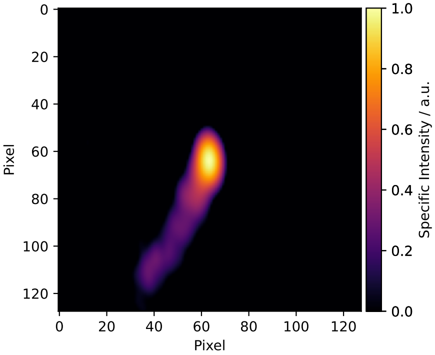


Figure: MOJAVE image of 1142+198 from 26.09.2016. [13]

Evaluation of MOJAVE data

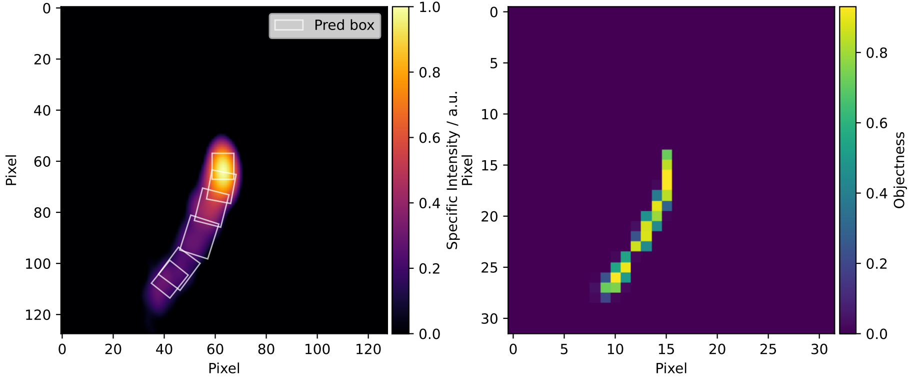


Figure: Evaluation of MOJAVE 1142+198 from 26.09.2016. [13]

Evaluation of MOJAVE data

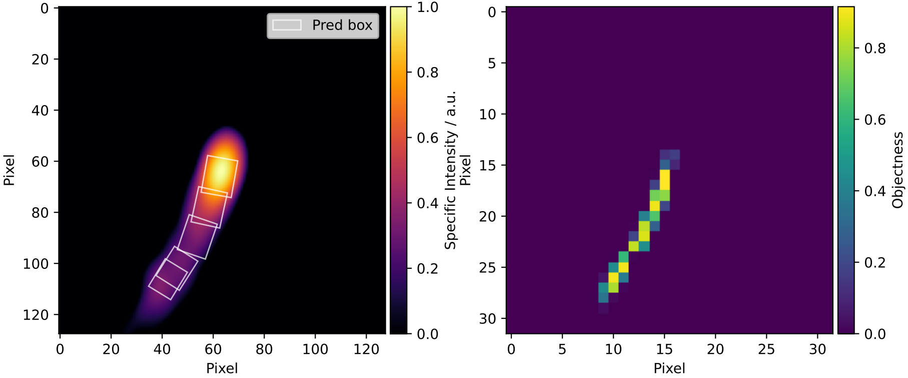


Figure: Evaluation of MOJAVE 1142+198 from 10.12.2016. [13]

Evaluation of MOJAVE data

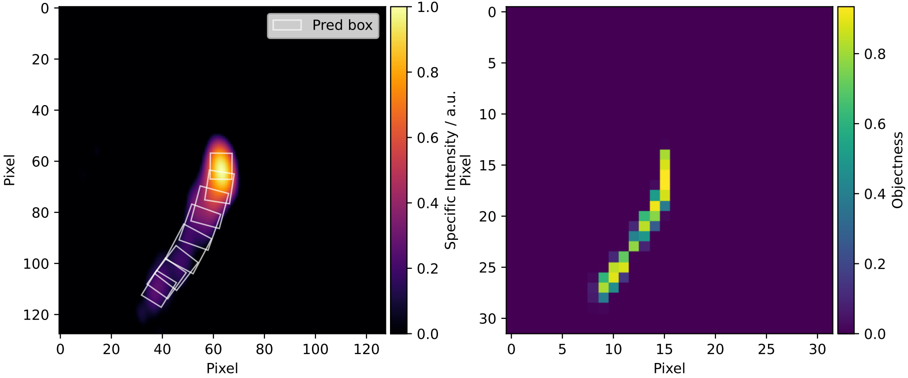


Figure: Evaluation of MOJAVE 1142+198 from 30.07.2017. [13]

Clustering

- Find similar components at each time step
- Spectral Clustering
  - Calculate affinity matrix
  - Scale with kNN distances
  - Reduce dimension by choosing largest eigenvalues
  - Eigenvectors create a low dimensional space
  - Perform kNN Clustering in lower dimension
- Build mean of equal components in one image

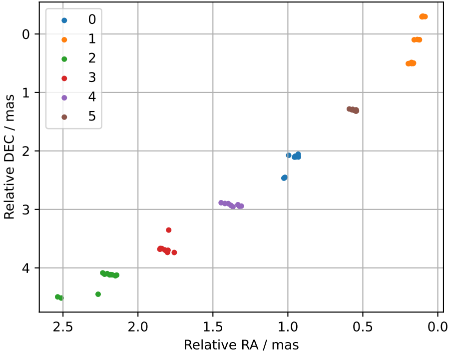


Figure: Spectral Clustering to group components.

Boxes after clustering

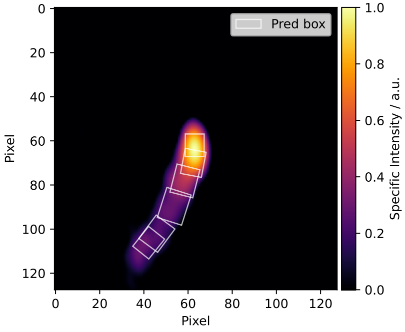


Figure: MOJAVE 1142+198 from 26.09.2016. [13]

Clustering

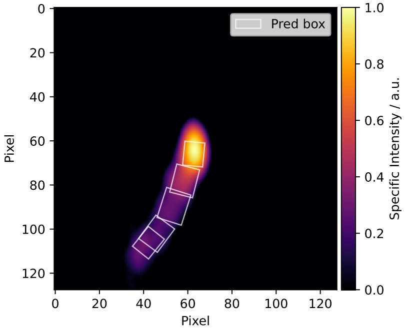


Figure: MOJAVE 1142+198 from 26.09.2016. [13]

Boxes after clustering

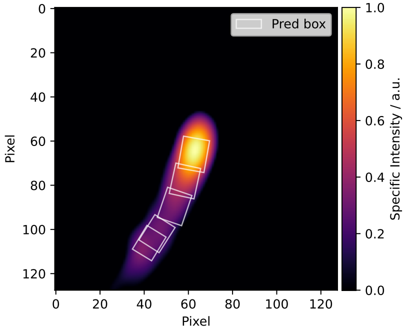


Figure: MOJAVE 1142+198 from 10.12.2016. [13]

Clustering →

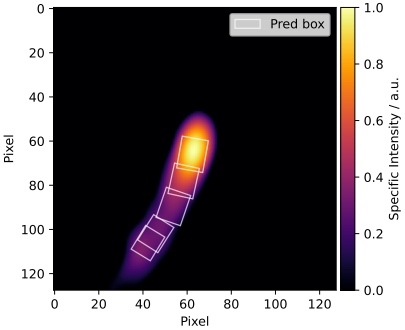


Figure: MOJAVE 1142+198 from 10.12.2016. [13]

Boxes after clustering

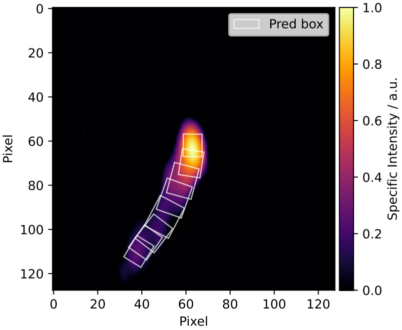


Figure: MOJAVE 1142+198 from 30.07.2017. [13]

Clustering →

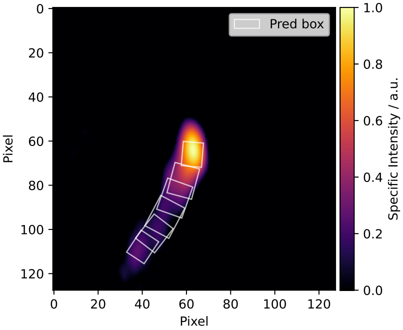


Figure: MOJAVE 1142+198 from 30.07.2017. [13]

Velocity calculation

- Components are located and assigned to each other
- Velocity in mas/yr from linear regression
- Converted into units of c_0 with given distance to source
- Performed by Kevin Schmidt manually with `DIFMAP` in 2018

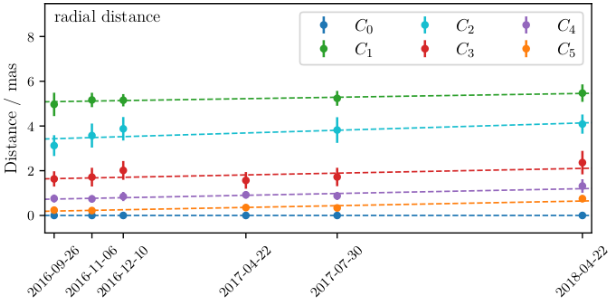


Figure: Kinematic analysis of 1142+198 by Kevin Schmidt. [14]

Velocity calculation

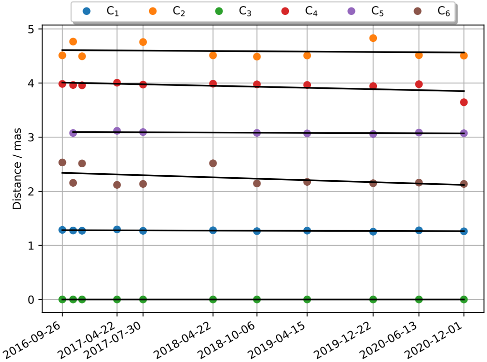


Figure: Distance of components and linear regression.

Component	Velocity / c_0
1	$-0{,}0068 \pm 0{,}0034$
2	$-0{,}0166 \pm 0{,}0485$
3	0
4	$-0{,}0590 \pm 0{,}0284$
5	$-0{,}0098 \pm 0{,}0062$
6	$-0{,}0838 \pm 0{,}0536$

Maximum jet speed by Lister et al. 2021 [15]:

$$(0{,}125 \pm 0{,}051)\,c$$

Conclusion

- Prediction of components with YOLO possible
- NMS critical part of analysis
- Clustering works as expected
- **But:** Velocity of components all negative

Component	Velocity / c_0
1	$-0{,}0068 \pm 0{,}0034$
2	$-0{,}0166 \pm 0{,}0485$
3	0
4	$-0{,}0590 \pm 0{,}0284$
5	$-0{,}0098 \pm 0{,}0062$
6	$-0{,}0838 \pm 0{,}0536$

Outlook

- For this work:
  - Reasonable performance metric(s)
  - Apply common methods on simulation for comparison
  - Analysis on more MOJAVE data

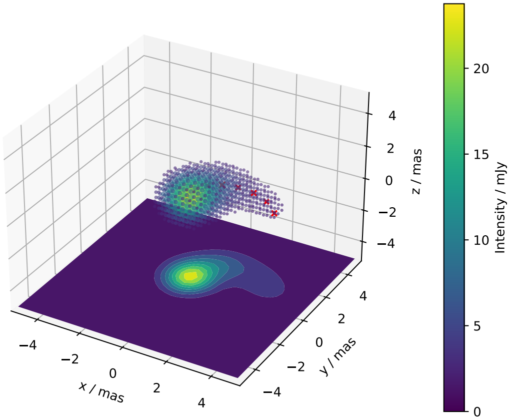


Figure: First approach of a realistic 3D simulation.

Outlook

- For this work:
  - Reasonable performance metric(s)
  - Apply common methods on simulation for comparison
  - Analysis on more MOJAVE data

- For future projects:
  - Find different architectures for reconstruction
  - Realistic simulations (physical properties, 3D, time development, ...)

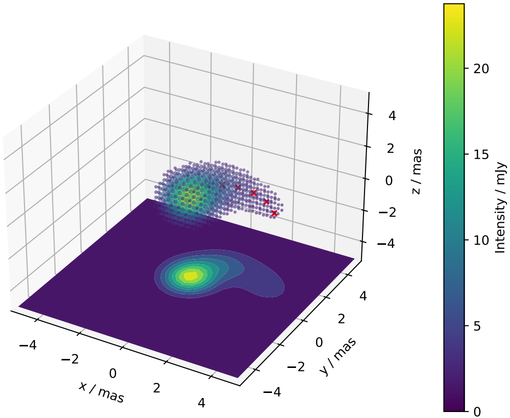


Figure: First approach of a realistic 3D simulation.

[1] Juan Antonio Aguilar and Jamie Yang. Cosmic messengers. IceCube/WIPAC.

[2] An Introduction to Radio Astronomy. URL: https://www.cv.nrao.edu/~sransom/web/Ch1.html (visited on 11/30/2022).

[3] Felix Geyer. "Reconstructing Radio Interferometric Data Using Neural Networks." Master of Science. TU Dortmund University, 2020.

[4] The MOJAVE Program Homepage. URL: https://www.cv.nrao.edu/MOJAVE/ (visited on 11/30/2022).

[5] Convolution Operator. URL: https://tikz.net/conv2d/ (visited on 11/30/2022).

[6] Y. V. R. Nagapawan, Kolla Bhanu Prakash, and G. R. Kanagachidambaresan. "Convolutional Neural Network." In: Programming with TensorFlow: Solution for Edge Computing Applications. Ed. by Kolla Bhanu Prakash and G. R. Kanagachidambaresan. Cham: Springer International Publishing, 2021, pp. 45–51. ISBN: 978-3-030-57077-4. DOI: 10.1007/978-3-030-57077-4_6. URL: https://doi.org/10.1007/978-3-030-57077-4_6.

[7] Joseph Redmon et al. "You Only Look Once: Unified, Real-Time Object Detection." In: CoRR abs/1506.02640 (2015). arXiv: 1506.02640. URL: http://arxiv.org/abs/1506.02640.

[8] Chuyi Li et al. YOLOv6: A Single-Stage Object Detection Framework for Industrial Applications. 2022. DOI: 10.48550/ARXIV.2209.02976. URL: https://arxiv.org/abs/2209.02976.

[9] Zhaohui Zheng et al. Distance-IoU Loss: Faster and Better Learning for Bounding Box Regression. 2019. DOI: 10.48550/ARXIV.1911.08287. URL: https://arxiv.org/abs/1911.08287.

[10] Kevin Schmidt et al. radionets: Imaging Radio Interferometric Data with Neural Networks. Version v1.0.0. Nov. 2020. URL: https://github.com/radionets-project/radionets.

[11] Paul-Simon Blomenkamp. "Automated Source Detection in Radio Interferometric Images using Computer Vision." Master of Science. TU Dortmund University, 2021.

[12] Schmidt Kevin, Geyer Felix, and Poggenpohl Arne. radiosim: Simulation of radio skies to create astrophysical data sets. URL: https://github.com/radionets-project/radiosim.

[13] MOJAVE 1142+198. URL: https://www.cv.nrao.edu/MOJAVE/sourcepages/1142+198.shtml (visited on 11/30/2022).

[14] Kevin Schmidt. "Study of Jet Characteristics of TeV Radio Galaxy Candidates based on VLBA and MAGIC Observations." Master of Science. TU Dortmund University, 2018.

[15] M. L. Lister et al. "Monitoring Of Jets in Active Galactic Nuclei with VLBA Experiments. XVIII. Kinematics and Inner Jet Evolution of Bright Radio-loud Active Galaxies." In: The Astrophysical Journal (2021). DOI: 10.3847/1538-4357/ac230f.