# Mapping global dynamics of benchmark creation and saturation in artificial intelligence

**Adriano Barbosa-Silva[1,*], Simon Ott[1,*], Kathrin Blagec[1], Jan Brauner[2,3] and Matthias Samwald[1,§]**

[1] Institute of Artificial Intelligence, Medical University of Vienna. Währingerstraße 25a, 1090, Vienna, Austria.

[2] Oxford Applied and Theoretical Machine Learning (OATML) Group, Department of Computer Science, University of Oxford, Oxford, UK.

[3] Future of Humanity Institute, University of Oxford, Oxford, UK.

[*] *Equal contribution*

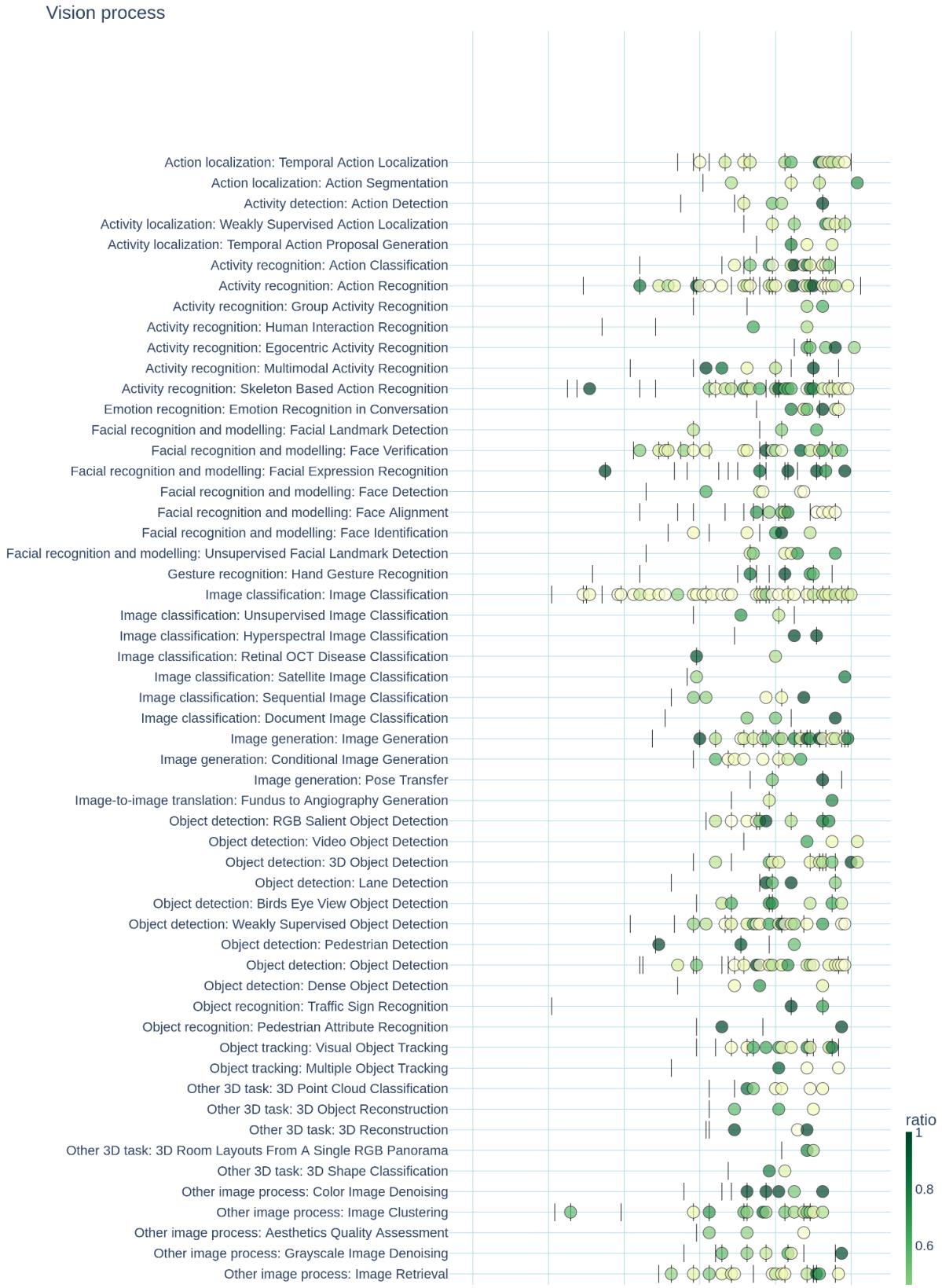
[§] *Corresponding author. matthias.samwald (at) meduniwien.ac.at*

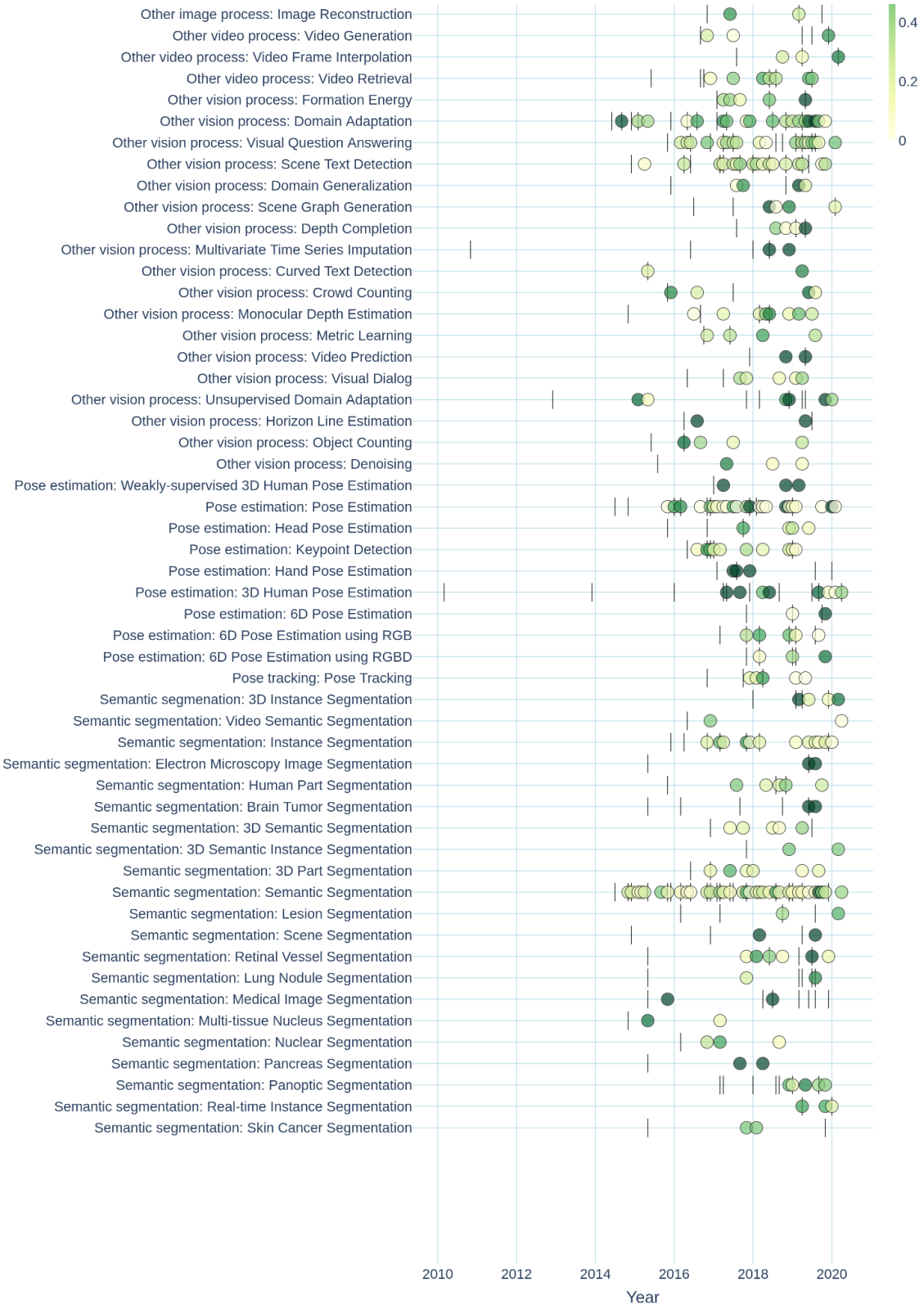
## Supplementary Information

**Suppl. Figure 1**: Global SOTA improvement map for computer vision. Anchors (vertical dashes) and only multiple dot trajectories shown in the plot.
**Suppl. Figure 2**: Global SOTA improvement map with single gain result for NLP.
**Suppl. Figure 3**: Global SOTA improvement map with single gain result for computer vision.
**Suppl. Figure 4**: Global SOTA improvement map with aggregated superclasses for NLP.
**Suppl. Figure 5**: Global SOTA improvement map with aggregated superclasses for computer vision.
**Suppl. Figure 6**: AI benchmark lifecycle map for computer vision.
**Suppl. Figure 7:** Number of active benchmarks vs. number of benchmarks reporting novel SOTA results over time for computer vision tasks.

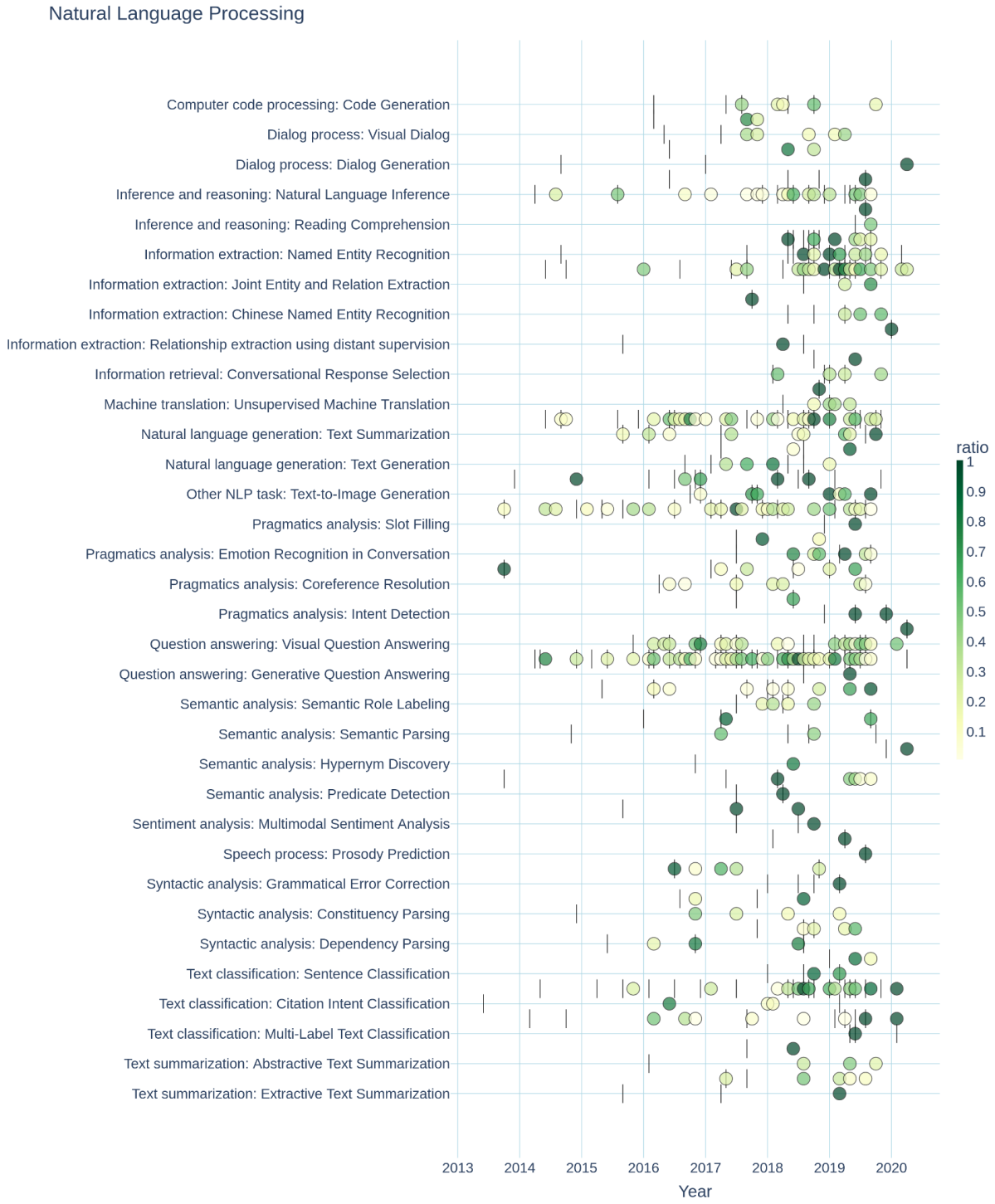
High definition images of all figures are available here:
https://drive.google.com/drive/folders/1OlU13253wzxVIbEMGHcB_duGIoHlVCO1?usp=sharing

**Supplementary Figure 1:** Global SOTA improvement map for computer vision. Anchors (vertical dashes) and only multiple dot trajectories shown in the plot.

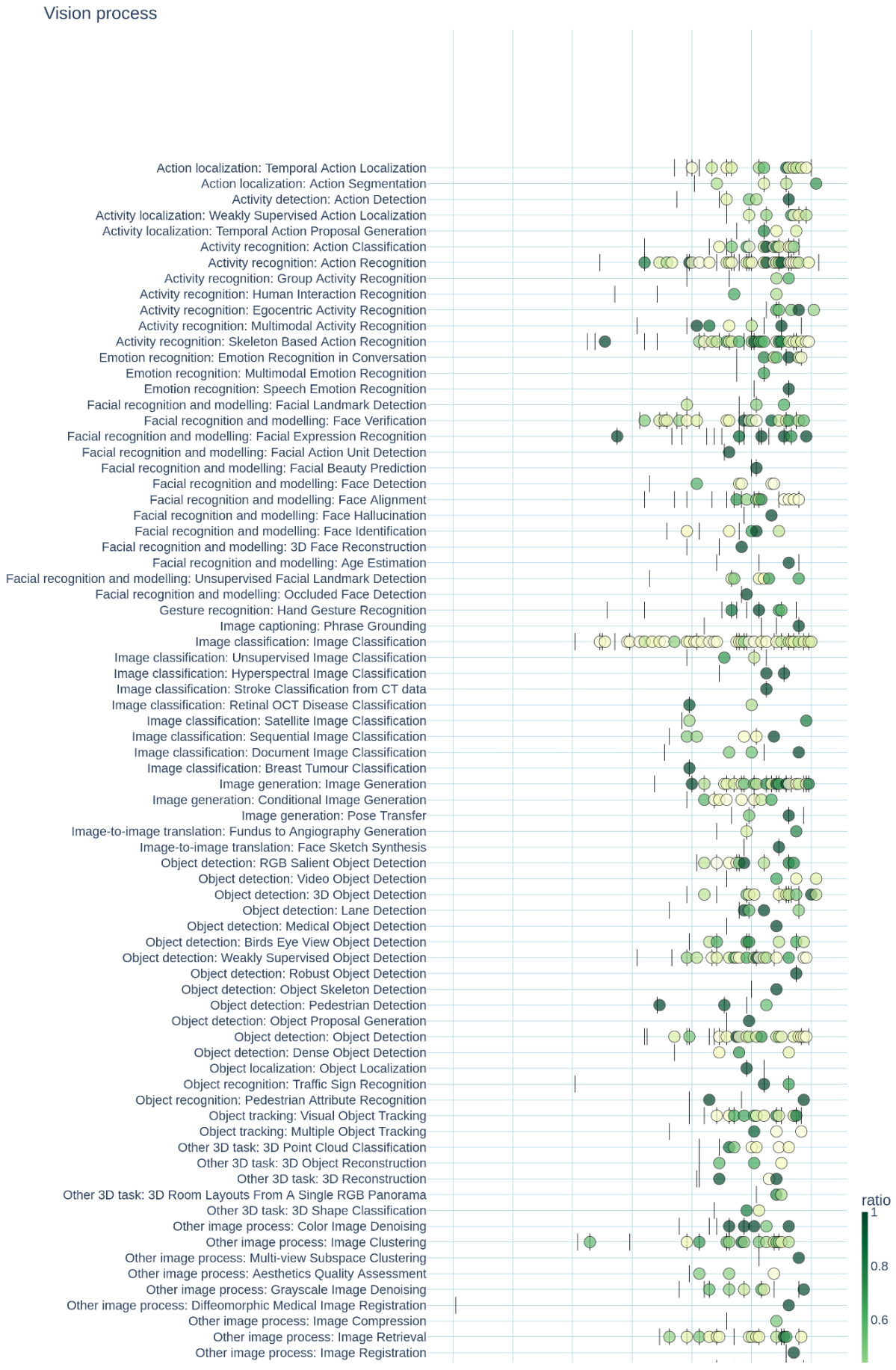


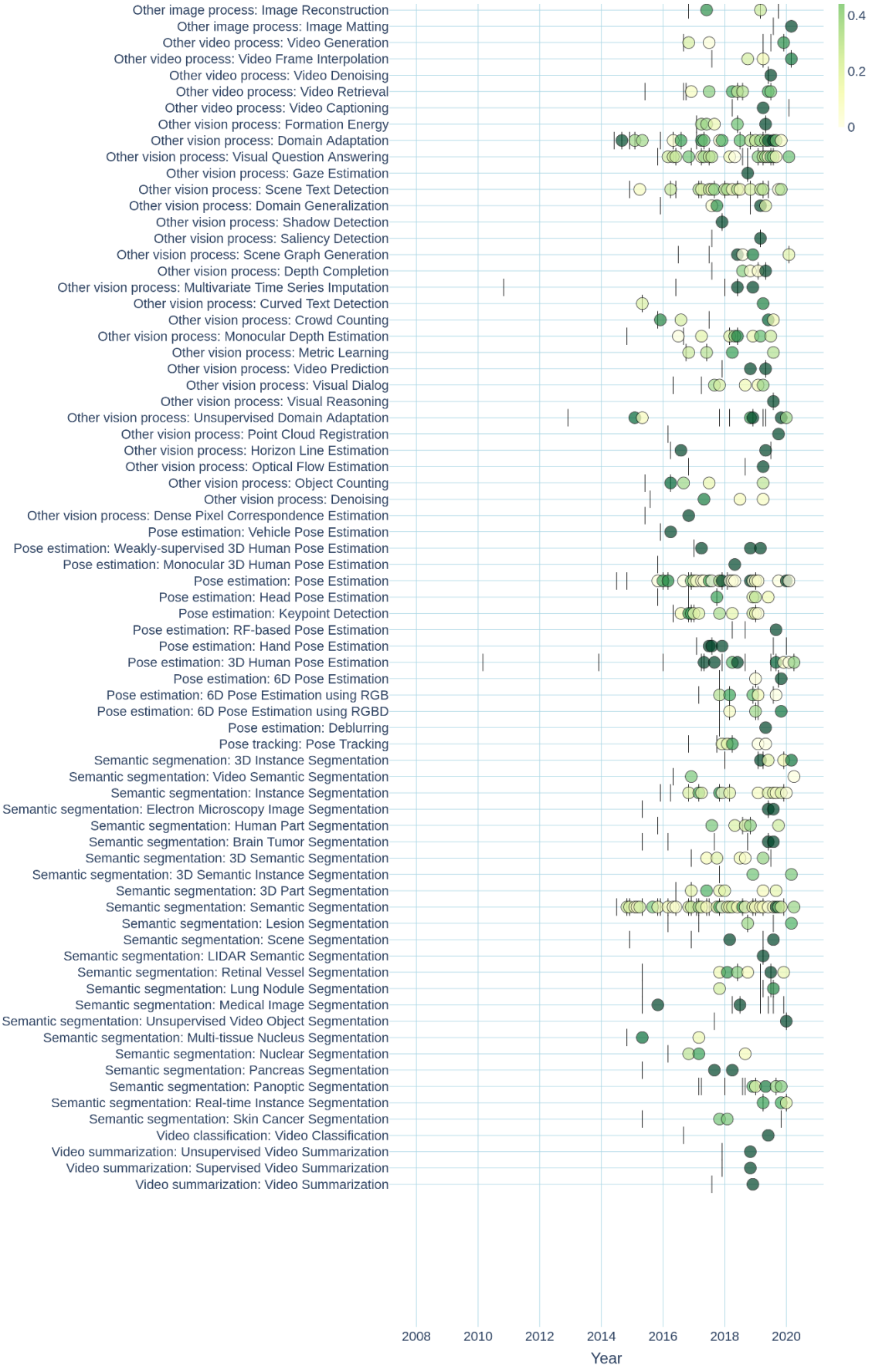

Other image process: Image Reconstruction
Other video process: Video Generation
Other video process: Video Frame Interpolation
Other video process: Video Retrieval
Other vision process: Formation Energy
Other vision process: Domain Adaptation
Other vision process: Visual Question Answering
Other vision process: Scene Text Detection
Other vision process: Domain Generalization
Other vision process: Scene Graph Generation
Other vision process: Depth Completion
Other vision process: Multivariate Time Series Imputation
Other vision process: Curved Text Detection
Other vision process: Crowd Counting
Other vision process: Monocular Depth Estimation
Other vision process: Metric Learning
Other vision process: Video Prediction
Other vision process: Visual Dialog
Other vision process: Unsupervised Domain Adaptation
Other vision process: Horizon Line Estimation
Other vision process: Object Counting
Other vision process: Denoising
Pose estimation: Weakly-supervised 3D Human Pose Estimation
Pose estimation: Pose Estimation
Pose estimation: Head Pose Estimation
Pose estimation: Keypoint Detection
Pose estimation: Hand Pose Estimation
Pose estimation: 3D Human Pose Estimation
Pose estimation: 6D Pose Estimation
Pose estimation: 6D Pose Estimation using RGB
Pose estimation: 6D Pose Estimation using RGBD
Pose tracking: Pose Tracking
Semantic segmenation: 3D Instance Segmentation
Semantic segmentation: Video Semantic Segmentation
Semantic segmentation: Instance Segmentation
Semantic segmentation: Electron Microscopy Image Segmentation
Semantic segmentation: Human Part Segmentation
Semantic segmentation: Brain Tumor Segmentation
Semantic segmentation: 3D Semantic Segmentation
Semantic segmentation: 3D Semantic Instance Segmentation
Semantic segmentation: 3D Part Segmentation
Semantic segmentation: Semantic Segmentation
Semantic segmentation: Lesion Segmentation
Semantic segmentation: Scene Segmentation
Semantic segmentation: Retinal Vessel Segmentation
Semantic segmentation: Lung Nodule Segmentation
Semantic segmentation: Medical Image Segmentation
Semantic segmentation: Multi-tissue Nucleus Segmentation
Semantic segmentation: Nuclear Segmentation
Semantic segmentation: Pancreas Segmentation
Semantic segmentation: Panoptic Segmentation
Semantic segmentation: Real-time Instance Segmentation
Semantic segmentation: Skin Cancer Segmentation
0.4
0.2
0
2010
2012
2014
2016
2018
2020
Year

**Suppl. Figure 2**: Global SOTA improvement map with single gain result for NLP.

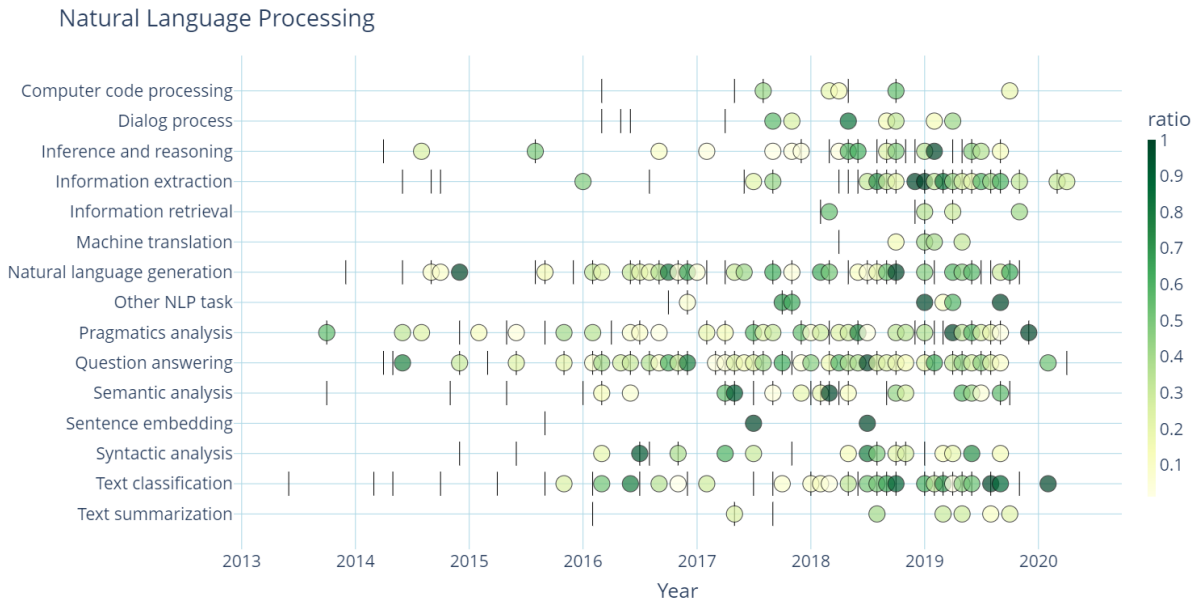


**Suppl. Figure 3**: Global SOTA improvement map with single gain result for computer vision.

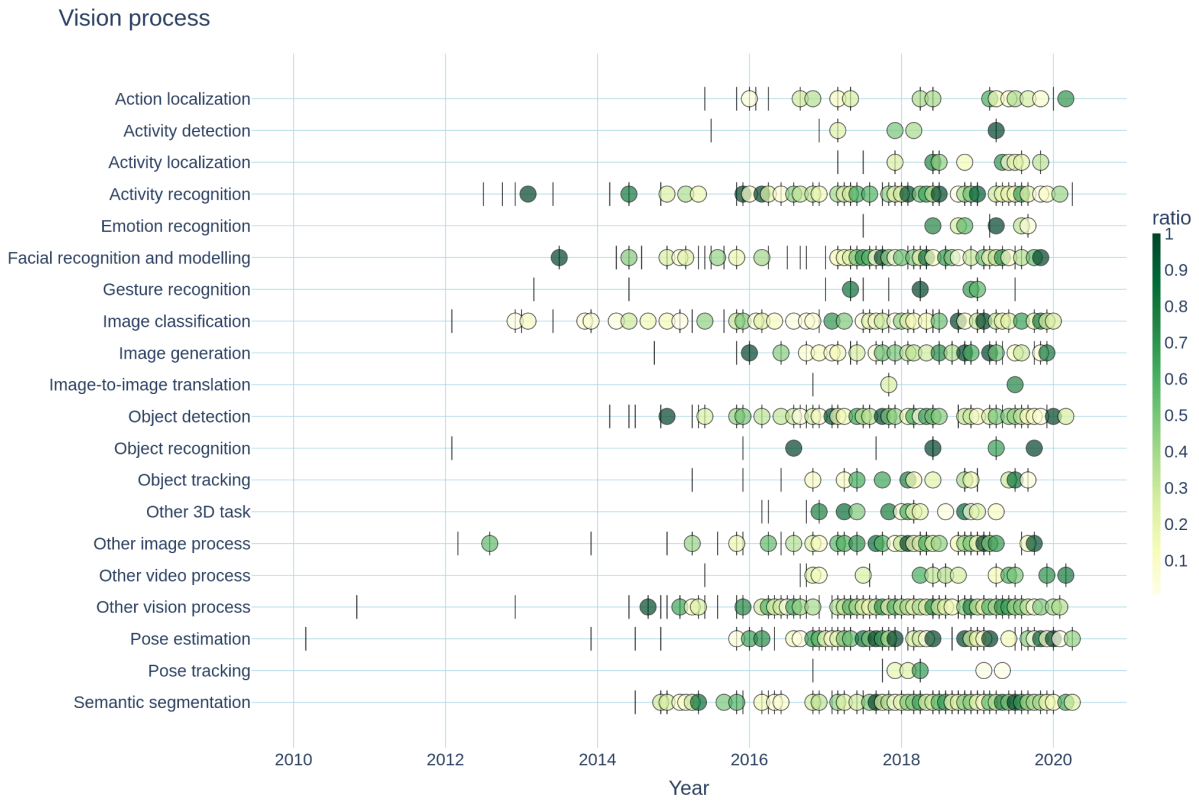



Other image process: Image Reconstruction
Other image process: Image Matting
Other video process: Video Generation
Other video process: Video Frame Interpolation
Other video process: Video Denoising
Other video process: Video Retrieval
Other video process: Video Captioning
Other vision process: Formation Energy
Other vision process: Domain Adaptation
Other vision process: Visual Question Answering
Other vision process: Gaze Estimation
Other vision process: Scene Text Detection
Other vision process: Domain Generalization
Other vision process: Shadow Detection
Other vision process: Saliency Detection
Other vision process: Scene Graph Generation
Other vision process: Depth Completion
Other vision process: Multivariate Time Series Imputation
Other vision process: Curved Text Detection
Other vision process: Crowd Counting
Other vision process: Monocular Depth Estimation
Other vision process: Metric Learning
Other vision process: Video Prediction
Other vision process: Visual Dialog
Other vision process: Visual Reasoning
Other vision process: Unsupervised Domain Adaptation
Other vision process: Point Cloud Registration
Other vision process: Horizon Line Estimation
Other vision process: Optical Flow Estimation
Other vision process: Object Counting
Other vision process: Denoising
Other vision process: Dense Pixel Correspondence Estimation
Pose estimation: Vehicle Pose Estimation
Pose estimation: Weakly-supervised 3D Human Pose Estimation
Pose estimation: Monocular 3D Human Pose Estimation
Pose estimation: Pose Estimation
Pose estimation: Head Pose Estimation
Pose estimation: Keypoint Detection
Pose estimation: RF-based Pose Estimation
Pose estimation: Hand Pose Estimation
Pose estimation: 3D Human Pose Estimation
Pose estimation: 6D Pose Estimation
Pose estimation: 6D Pose Estimation using RGB
Pose estimation: 6D Pose Estimation using RGBD
Pose estimation: Deblurring
Pose tracking: Pose Tracking
Semantic segmenation: 3D Instance Segmentation
Semantic segmentation: Video Semantic Segmentation
Semantic segmentation: Instance Segmentation
Semantic segmentation: Electron Microscopy Image Segmentation
Semantic segmentation: Human Part Segmentation
Semantic segmentation: Brain Tumor Segmentation
Semantic segmentation: 3D Semantic Segmentation
Semantic segmentation: 3D Semantic Instance Segmentation
Semantic segmentation: 3D Part Segmentation
Semantic segmentation: Semantic Segmentation
Semantic segmentation: Lesion Segmentation
Semantic segmentation: Scene Segmentation
Semantic segmentation: LIDAR Semantic Segmentation
Semantic segmentation: Retinal Vessel Segmentation
Semantic segmentation: Lung Nodule Segmentation
Semantic segmentation: Medical Image Segmentation
Semantic segmentation: Unsupervised Video Object Segmentation
Semantic segmentation: Multi-tissue Nucleus Segmentation
Semantic segmentation: Nuclear Segmentation
Semantic segmentation: Pancreas Segmentation
Semantic segmentation: Panoptic Segmentation
Semantic segmentation: Real-time Instance Segmentation
Semantic segmentation: Skin Cancer Segmentation
Video classification: Video Classification
Video summarization: Unsupervised Video Summarization
Video summarization: Supervised Video Summarization
Video summarization: Video Summarization
0.4
0.2
0
2008
2010
2012
2014
2016
2018
2020
Year

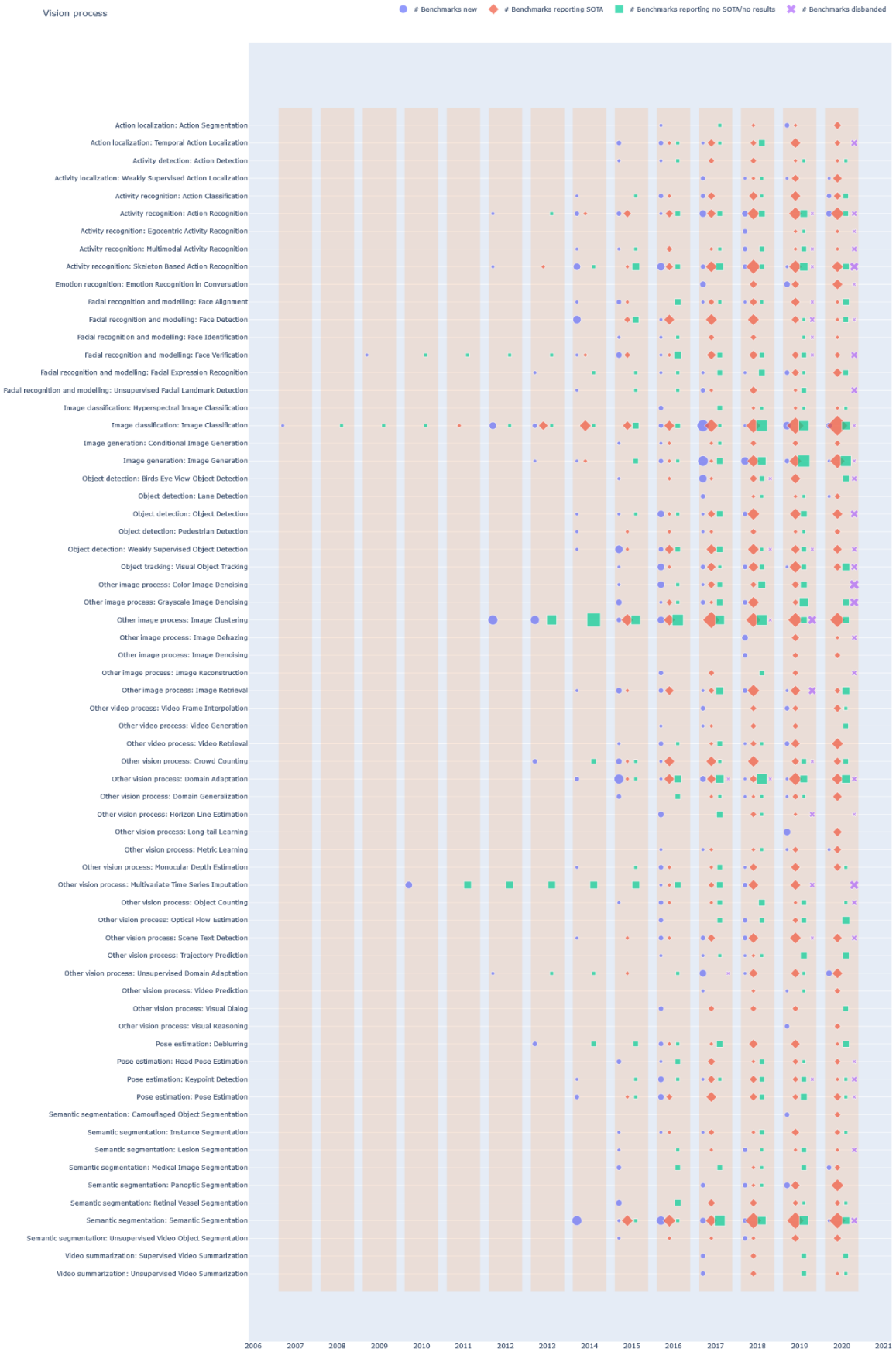
**Suppl. Figure 4**: Global SOTA improvement map with aggregated superclasses for NLP.




**Supplementary Figure 5:** *Global SOTA improvement map for computer vision with aggregated superclasses.*



**Supplementary Figure 6**: AI benchmark lifecycle map for computer vision.



**Supplementary Figure 7:** Number of active benchmarks vs. number of benchmarks reporting novel SOTA results over time for computer vision tasks.