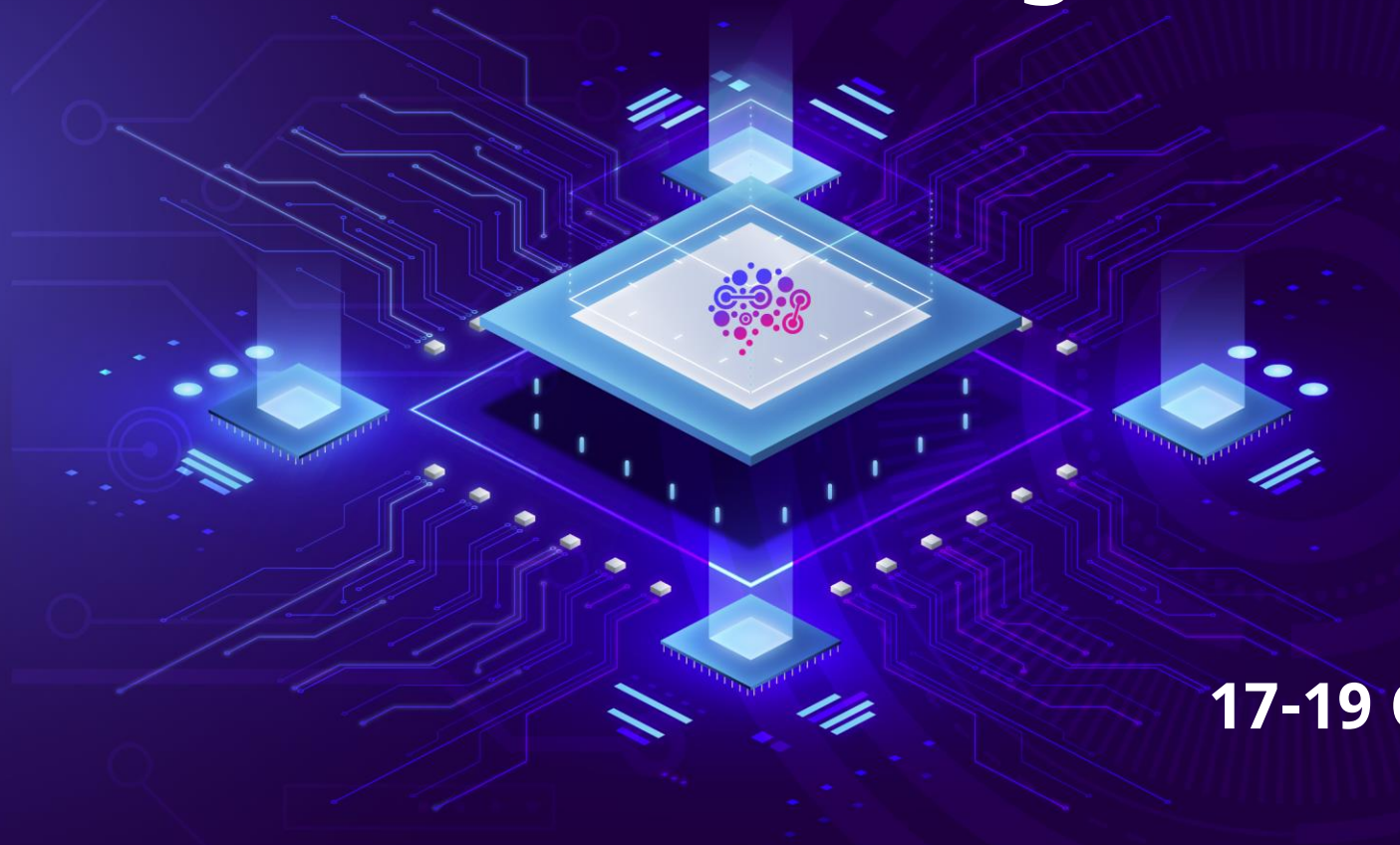
European Conference on EDGE AI Technologies and Applications - EEAI



Advancing emerging edge AI technologies and driving next generation intelligent applications.

17-19 October 2023, Athens, Greece

European Conference on EDGE AI Technologies and Applications - EEAI

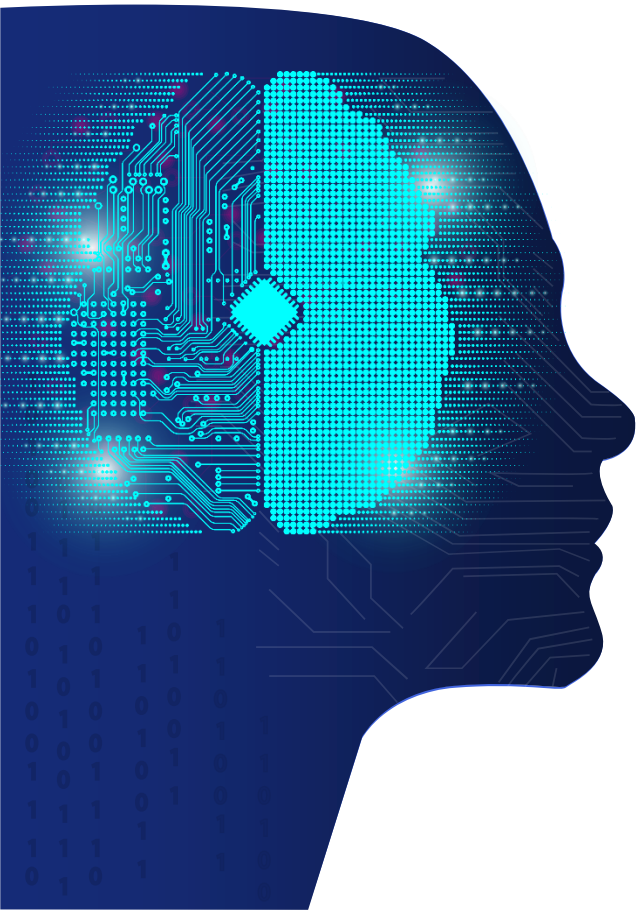
VC3 – Agrifood and Beverage VCD3.1

HOLLARD Lilian, URCA, France

17-19 October 2023 Athens, Greece

Copyright © 2023

Presentation Outline

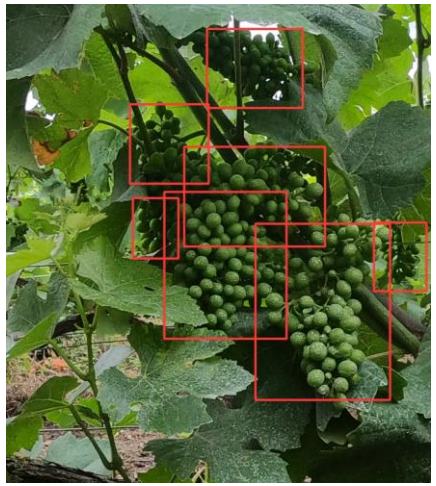


- Introduction
- State-of-the-art
- Research Goal
- Methodology
- Design and Implementation
- Experimental Results
- Future Research
- Conclusion

Copyright © 2023

Introduction

- **VCD 3.1**: Yield Estimation (Partners: **URCA**, VPHV, CV, STM-FR)
  - Provide decisive information for the organization of the harvest
    - Use RGB cameras and Edge device mounted in an autonomous robot (Vitibot® Bakus)
    - AI Inference at the Meta (GPU) Edge level for grape counting
    - Detecting grapes at BBCH77 phenological state (End of June)
      - High environment density (occlusion) and difficult to distinguish



Partners

UNIVERSITÉ DE REIMS CHAMPAGNE-ARDENNE

VRANKEN POMMERY MONOPOLE

life.augmented

Copyright © 2023

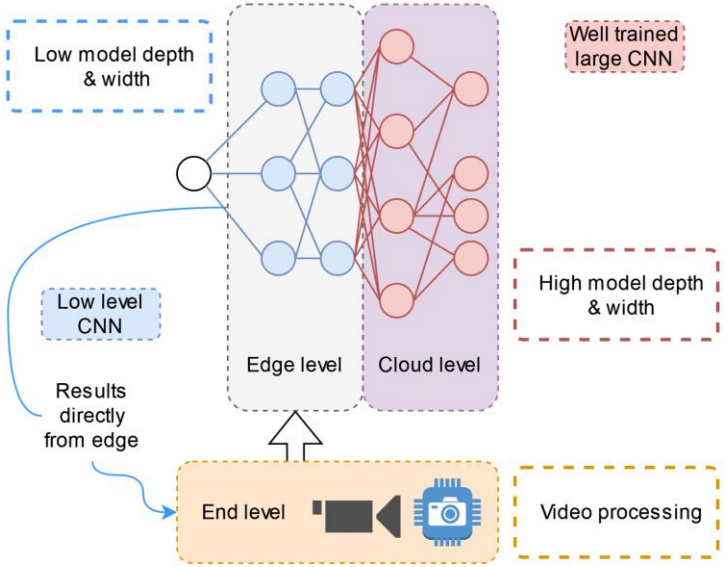
State-of-the-art

- Problematics
  - Yield forecast currently performed by manual sampling on the vineyard
    - High variability between plots: more than 30% error
  - Not enough variety of early grapevine dataset
    - WGISD Dataset – Grapevine dataset: Fully matured grape, leaves removed
    - Grape MS Dataset – Grapevine dataset: Early stage but not Champagne-related grapevine types
  - Deep Learning related solution :
    - Not mobile-oriented
    - No autonomous solution has yet been discovered

Copyright © 2023

Research Goal

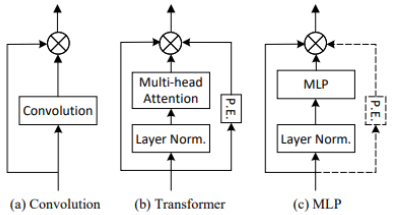
- Reach yield forecast below a 15% error margin
  - Develop AI-Edge based platforms for end-to-end SW/HW stack {01}
    - Edge AI application scenario
    - Object recognition accuracy with >85% mAP and additional correlation refinements
- Algorithm adaption to different Edge processing platforms
  - Scalable and modular AI into the novel HW/SW AI open architecture platforms {04}
    - STM32
    - GPUs
    - TPUs

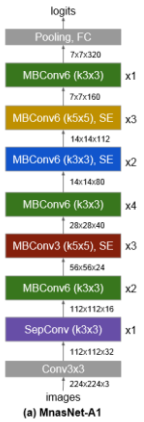


Copyright © 2023

Methodology

- Researching a vast spectrum of model architecture
  - Room for optimization
  - Find better training recipes
  - Investigate scalable architecture



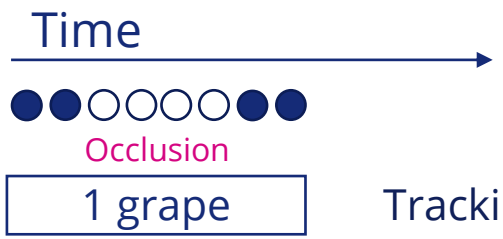


- Create our own dataset
  - Images collected at different stages (from flower buds to maturated grapes)

Copyright © 2023

Design and Implementation

- Object detection
  - Detecting grapes with a fast object detector
    - YOLO – Best accuracy/speed compromise
    - We want real-time video processing (at least 20 to 25 frames per second)
  - High processing framerate
    - Vehicle speed: 3-4 km/h
    - Camera location: 50-60cm away from the vines
    - On average, a vine passes through in 1 to 2 seconds max.
    - Better tracking possibilities
      - Kalman Filter (occlusion correction)

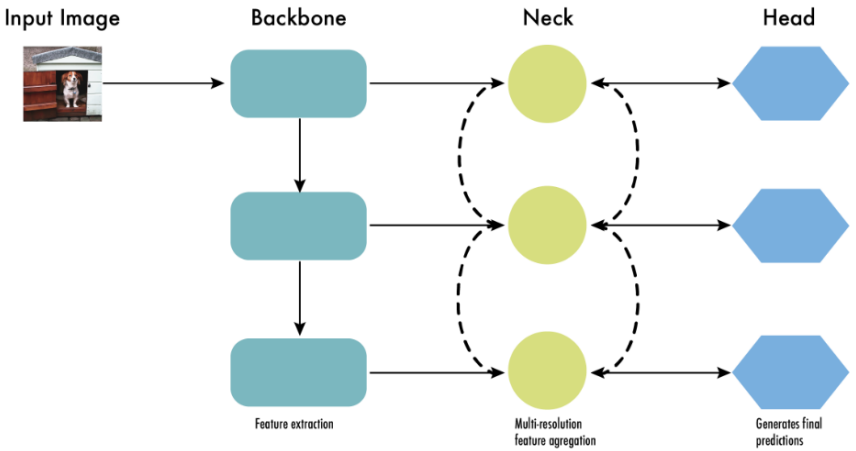


Tracking grapes over the X axis

Copyright © 2023

Design and Implementation

- Object detection
  - Detecting grapes with a fast object detector
    - YOLO – Best accuracy/speed compromise
    - We want real-time video processing (at least 20 to 25 frames per second)
  - Optimize YOLO for Edge devices
    - Scale model channels and depths
    - Optimize the semantic level of information



Do you need as much information pathway as for the MSCOCO Dataset ?

Copyright © 2023

Experimental Results

- Grapevines within 640p images are relatively small
  - BBCH77 (early grapevine) is not fully grown yet
  - 640p image processing is much more reasonable for high framerate processing
- Only ONE class to detect
  - One-class models are much more scalable than models designed for 80 to 1000 classes.

Area < 32^2

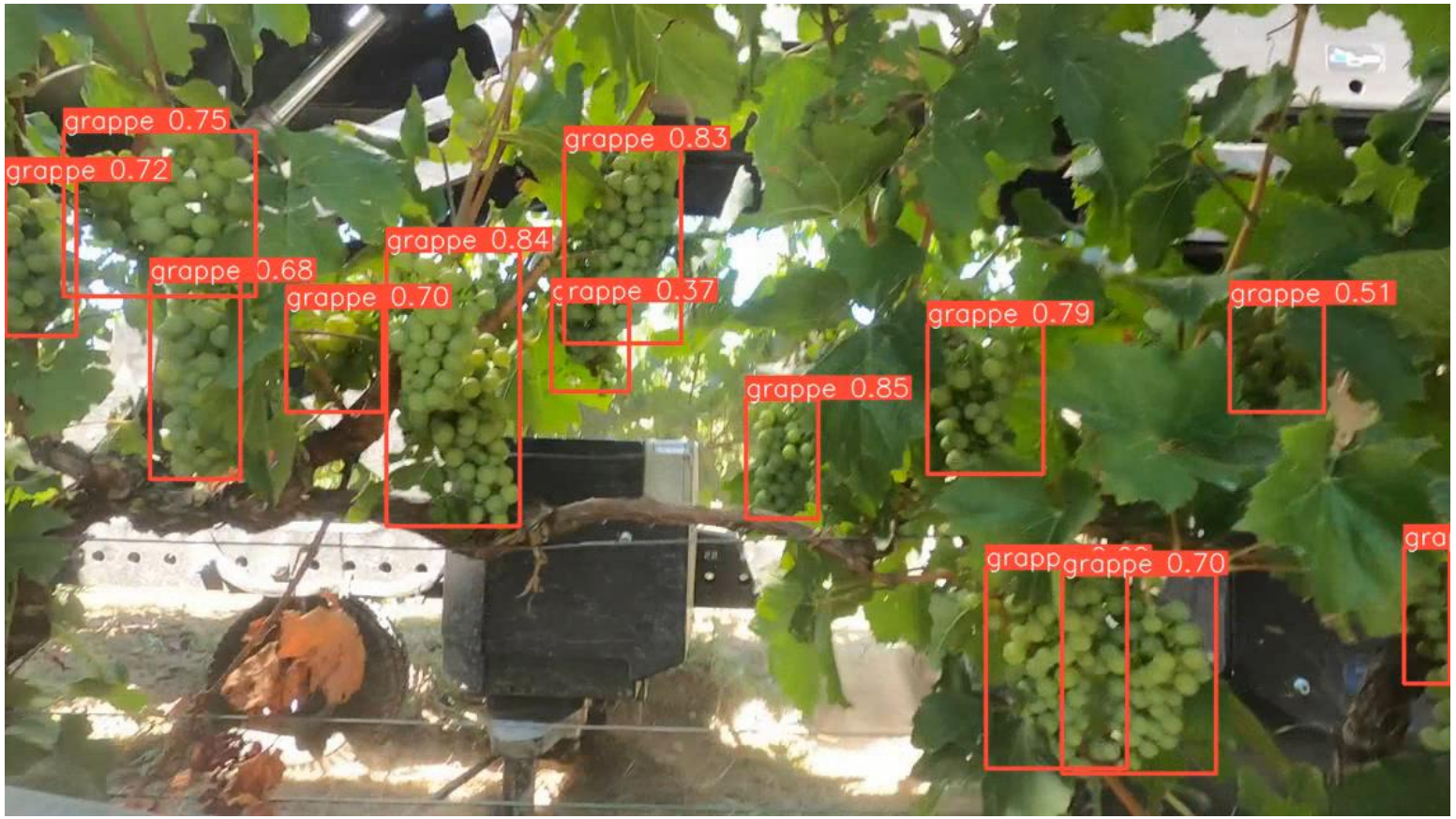
After 4 downsampling

1x1 pixel
if input = 640p

Copyright © 2023

Experimental Results

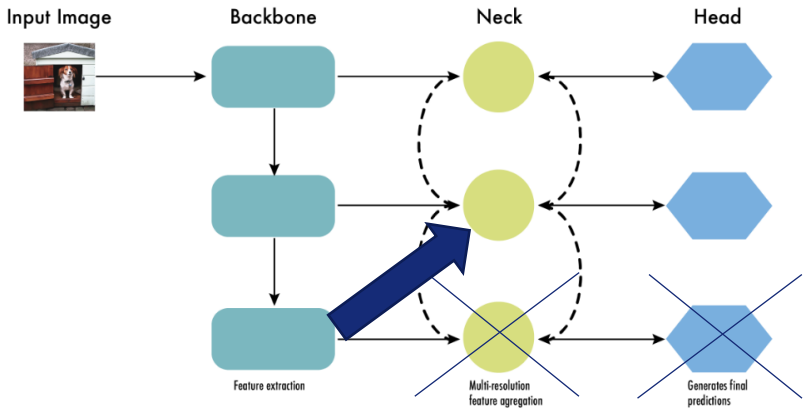
- Showcase
  - Experimental test



Copyright © 2023

Future Research

- Optimization of YOLO based model
  - Less activation map processing
    - Fewer channels
  - Optimized neck information transition pathway
  - Achieved super computed deep neural network accuracy with mobile or edge-optimized model architecture



- Optimized neck
- Fewer channels
- Less depth

Using foundation models to improve training recipes

Copyright © 2023

Future Research

- Optimization of grapes detection
  - Several architectures might be more interesting
    - MobileNetv3, HRNets, RepCOL, TinyViT, FastViT, MobileNeXt, MobileOne, Res2Net, …
  - Optimize the way detectors are working
    - Is there a more parallelizable way to construct a neck for object detection ?
- **Investigating the influence of occlusion**
  - 30% to 40% of grapes occlusion in 2022 (50 vines captured before and after leaf removal)
    - Study renewed in 2023 (180 vines captured)

Copyright © 2023

Future Research

- Optimization of grapes detection
  - Better training possibility (Segmentation oriented training)

Use the backbone of the model trained on segmented data

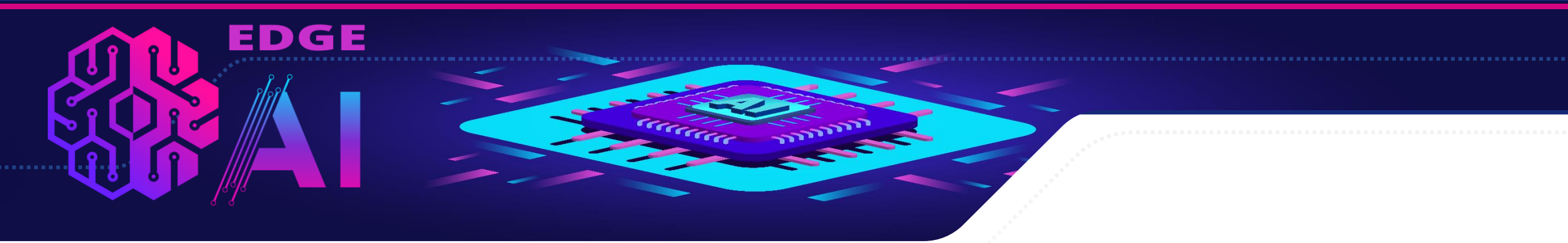
- Same backbone:
  - Yolov5 Nano (object detection) : 4,5G
  - Yolov5Seg Nano (segmentation): 7.1G

Copyright © 2023

Conclusion

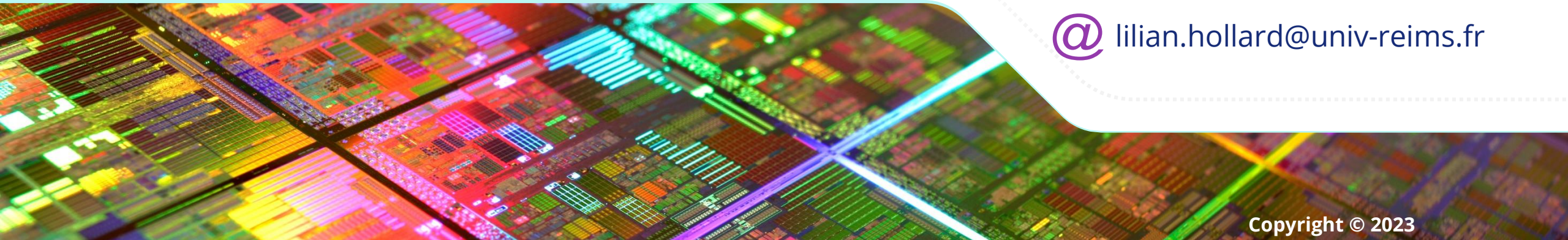
- The first step towards a grapevine yield forecast tool for Edge-AI embedded device
  - Improved optimized and scaled YOLO architecture
  - Working on our own object detection model for performant grapevine detection
  - Reducing the gap between the super computed oriented model and the mobile/edge-oriented model
- Development of a yield density map
  - RTK-GPS coupled with grape counting

Copyright © 2023



Thank You

For your attention



@ lilian.hollard@univ-reims.fr

Copyright © 2023