

A Comparative Assessment of Multi-view Fusion Learning for Crop Classification

IGARSS
19.07.2023

Francisco Mena (1,2), **Diego Arenas** (2), **Marlon Nuske** (2), **Andreas Dengel** (1,2)

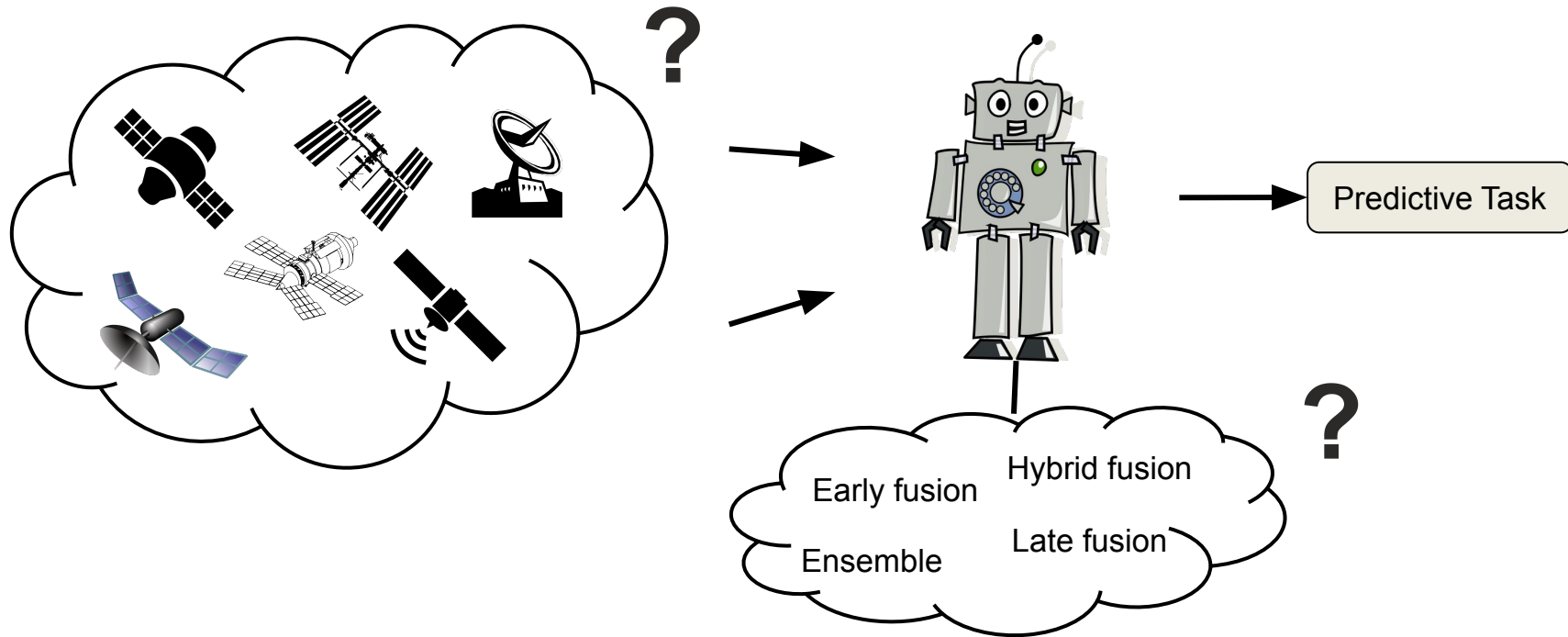
1 University of Kaiserslautern-Landau (RPTU)

2 German Research Center for Artificial Intelligence (DFKI)

Paper ID: 2532
email: menatoro@rptu.de

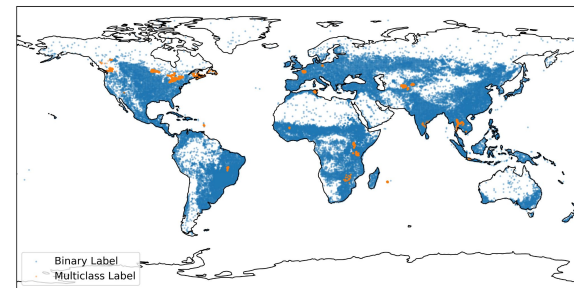


Context



Case Study

- **Crop Classification:** Is there a specific crop growing in an area during a season?
- Dataset: **CropHarvest** [1]
 - Pixel-wise **binary** classification task.
 - Two African countries as benchmark regions.
 - We include “Global” data.



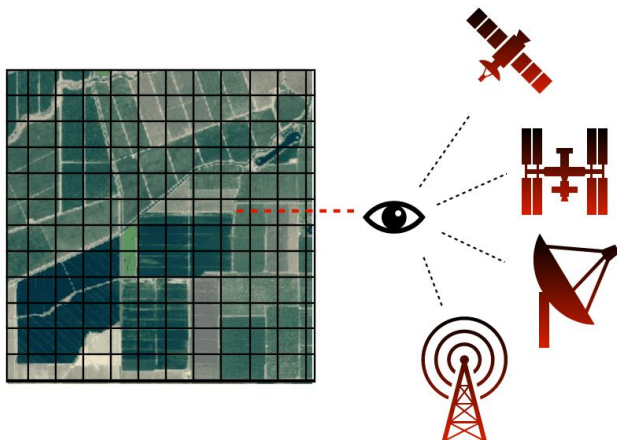
Data	Training	Testing
Kenya	1319 (20.0%)	898 (64.0%)
Togo	1290 (55.0%)	306 (34.6%)
Global	45723 (66.4%)	19520 (66.0%)

number of pixels (% of positive labeled)

[1] Tseng, Gabriel, et al. *Cropharvest: A global dataset for crop-type classification*. 2021

Case Study - Input data

- **Crop Classification:** Is there a specific crop growing in an area during a season?
- Dataset: **CropHarvest** [1]
 - 5 Multi-view (MV) input data:



- Optical (*from S2*):
 - 12 time-steps (monthly) x 11 bands
- NDVI (*from Optical*):
 - 12 time-steps x 1 band
- Radar (*from S1*):
 - 12 time-steps x 2 bands
- Weather (*from ERA5*):
 - 12 time-steps x 2 bands
- DEM (*from SRTM*):
 - vector with 2 bands

[1] Tseng, Gabriel, et al. *Cropharvest: A global dataset for crop-type classification*. 2021

Single-view Learning

- Evidence that **single-view learning** with some RS views does not achieve good classification.
 - Difference to conventional multi-modal learning (audio, text, image).

Data	View	AA
Kenya	DEM	48.0 $\pm$ 0.9
	Weather	50.0 $\pm$ 0.0
	Radar	63.0 $\pm$ 1.0
Togo	DEM	61.2 $\pm$ 2.6
	Weather	55.6 $\pm$ 3.5
	Optical	80.0 $\pm$ 1.3
Global	DEM	64.9 $\pm$ 0.1
	Weather	72.7 $\pm$ 0.8
	Optical	78.7 $\pm$ 0.8

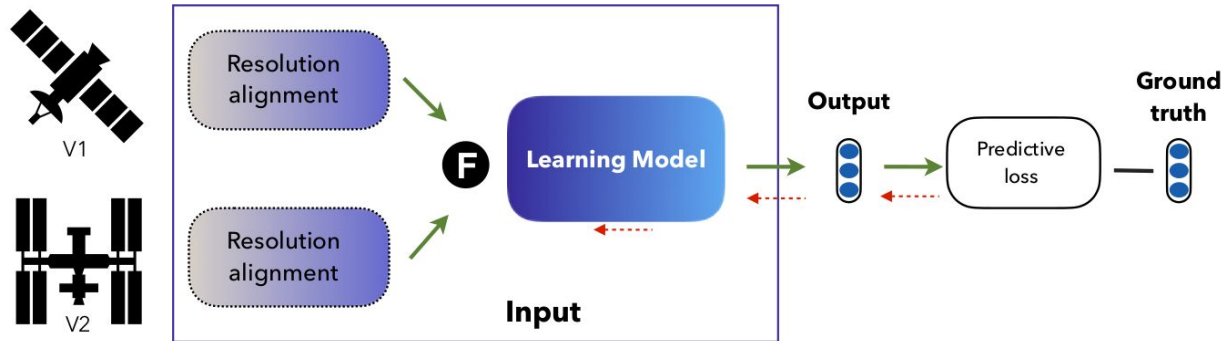
AA = Average Accuracy

In RS, each source could provide complementary information that needs to be compared and merged.

MV Learning: Input

Which is the best multi-view fusion approach for the case study?

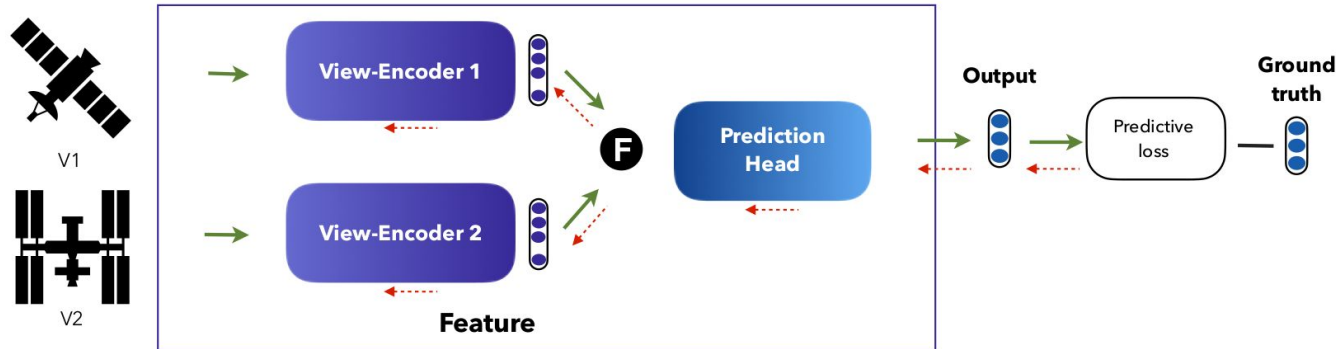
- **Input**, Feature, Decision, Ensemble



MV Learning: Feature

Which is the best multi-view fusion approach for the case study?

- Input, **Feature**, Decision, Ensemble
 1. Default: average/concat.
 2. Gated Fusion [2]: weighted sum.
 3. Multi Loss [3]: auxiliary classifiers.



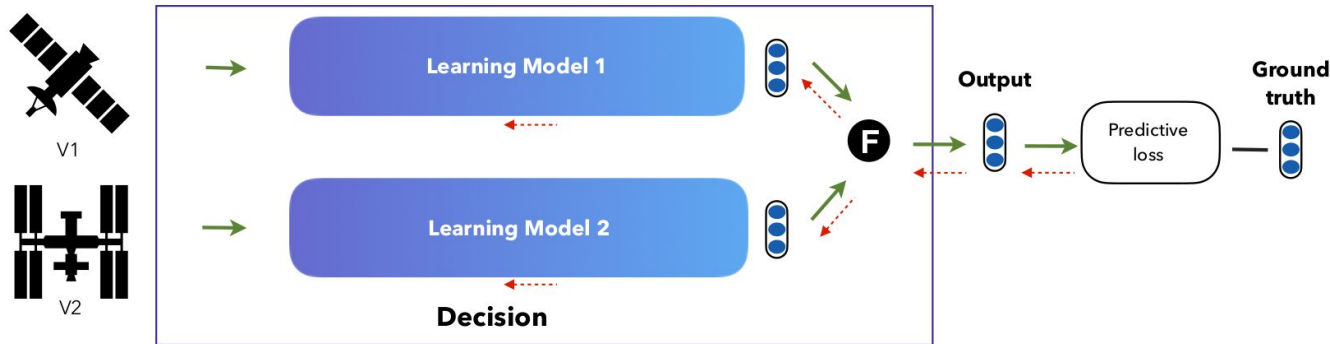
[2] Arevalo, John, et al. *Gated multimodal networks*. 2020.

[3] Benedetti, Paola, et al. *M3 Fusion: A Deep Learning Architecture for Multiscale Multimodal Multitemporal Satellite Data Fusion*. 2018.

MV Learning: Decision

Which is the best multi-view fusion approach for the case study?

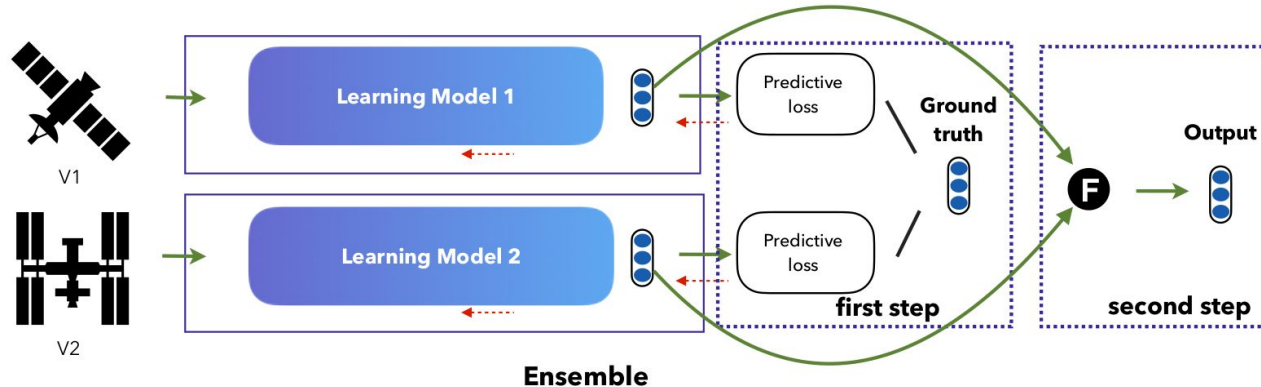
- Input, Feature, **Decision**, Ensemble



MV Learning: Ensemble

Which is the best multi-view fusion approach for the case study?

- Input, Feature, Decision, **Ensemble**



Results - Kenya, Togo

Data	Method	AA	AUC
Kenya	<i>Radar</i>	63.0 ± 1.0	66.8 ± 2.9
	Input	61.3 ± 7.5	70.0 ± 7.2
	Feature	63.0 ± 5.1	71.6 ± 3.7
	Gated Fusion	66.5 ± 3.6	71.8 ± 3.5
	Multi-Loss	64.8 ± 5.2	71.6 ± 4.1
	Decision	57.5 ± 7.6	63.2 ± 7.0
	Ensemble	56.0 ± 5.3	69.5 ± 1.4
Togo	<i>Optical</i>	80.0 ± 1.3	88.6 ± 0.7
	Input	79.7 ± 1.5	89.0 ± 1.3
	Feature	79.9 ± 1.0	88.9 ± 0.8
	Gated Fusion	78.1 ± 1.8	87.6 ± 1.2
	Multi-Loss	78.2 ± 6.1	88.6 ± 1.7
	Decision	81.5 ± 1.6	89.5 ± 0.5
	Ensemble	84.0 ± 1.0	90.9 ± 0.5

- The best performing method is not the **same across testing regions**.
- The **top 3 performance MV methods** improve single-view learning results.
 - Some decrease performance.

Metrics: Average Accuracy (AA), Area Under Curve (AUC)
Optimization: ADAM, 256 batch size, early stopping.

Results - Kenya, Togo

Data	Method	AA	AUC
Kenya	<i>Radar</i>	63.0 ± 1.0	66.8 ± 2.9
	Input	61.3 ± 7.5	70.0 ± 7.2
	Feature	63.0 ± 5.1	71.6 ± 3.7
	Gated Fusion	66.5 ± 3.6	71.8 ± 3.5
	Multi-Loss	64.8 ± 5.2	71.6 ± 4.1
	Decision	57.5 ± 7.6	63.2 ± 7.0
	Ensemble	56.0 ± 5.3	69.5 ± 1.4
Togo	<i>Optical</i>	80.0 ± 1.3	88.6 ± 0.7
	Input	79.7 ± 1.5	89.0 ± 1.3
	Feature	79.9 ± 1.0	88.9 ± 0.8
	Gated Fusion	78.1 ± 1.8	87.6 ± 1.2
	Multi-Loss	78.2 ± 6.1	88.6 ± 1.7
	Decision	81.5 ± 1.6	89.5 ± 0.5
	Ensemble	84.0 ± 1.0	90.9 ± 0.5

- The best performing method is not the **same across testing regions**.
- The **top 3 performance MV methods** improve single-view learning results.
 - Some decrease performance.
- AUC result in the benchmark with RF [1]:
 - Kenya: **57.8 ± 0.6**
 - Togo: **89.2 ± 0.1**

[1] Tseng, Gabriel, et al. *Cropharvest: A global dataset for crop-type classification*. 2021

Results - Global

Data	Method	AA	AUC
Global	<i>Optical</i>	78.7 ± 0.8	86.7 ± 0.8
	Input	81.2 ± 0.9	89.7 ± 0.8
	Feature	81.3 ± 1.1	89.6 ± 1.1
	Gated Fusion	82.1 ± 1.4	90.6 ± 1.1
	Multi-Loss	81.6 ± 1.4	90.1 ± 1.1
	Decision	79.1 ± 1.4	87.8 ± 1.3
	Ensemble	79.2 ± 0.8	87.5 ± 1.5

- Global data has a similar behavior than Kenya data.
 - All MV fusion methods increase single-view learning results.

Results - Entropy

Data	Method	AA	AUC	Entropy
Kenya	<i>Radar</i>	63.0 ± 1.0	66.8 ± 2.9	69.2 ± 3.8
	Input	61.3 ± 7.5	70.0 ± 7.2	72.9 ± 6.1
	Feature	63.0 ± 5.1	71.6 ± 3.7	73.9 ± 4.4
	Gated Fusion	66.5 ± 3.6	71.8 ± 3.5	71.6 ± 5.8
	Multi-Loss	64.8 ± 5.2	71.6 ± 4.1	73.4 ± 4.3
	Decision	57.5 ± 7.6	63.2 ± 7.0	76.8 ± 7.6
	Ensemble	56.0 ± 5.3	69.5 ± 1.4	83.5 ± 2.2
Togo	<i>Optical</i>	80.0 ± 1.3	88.6 ± 0.7	62.7 ± 2.1
	Input	79.7 ± 1.5	89.0 ± 1.3	54.3 ± 4.2
	Feature	79.9 ± 1.0	88.9 ± 0.8	57.6 ± 4.1
	Gated Fusion	78.1 ± 1.8	87.6 ± 1.2	52.1 ± 7.8
	Multi-Loss	78.2 ± 6.1	88.6 ± 1.7	60.4 ± 12.2
	Decision	81.5 ± 1.6	89.5 ± 0.5	63.7 ± 2.5
	Ensemble	84.0 ± 1.0	90.9 ± 0.5	93.5 ± 0.6

$$\mathbb{H} = -P(\text{crop})\log P(\text{crop}) - P(\neg \text{crop})\log P(\neg \text{crop})$$

- Gated Fusion decreases the most the entropy of the predicted probabilities between MV methods.
 - Ensemble increases the most.
- **Merging towards the output-level the entropy tends to increase.**

↑ Entropy → ↑ prediction uncertainty
 ↓ Entropy → ↓ prediction uncertainty

Conclusions

- Improvements could be obtained by selecting an appropriate fusion strategy depending on the region.
 - One size does not fit all.
 - We achieve state-of-the-art performance in the African testing regions **without pre-training**.
- What is the most appropriate fusion approach for each case?
 - Preliminary criterion:
 - Positive biased labels: Feature-level fusion.
 - Negative biased labels: Decision-level fusion.
- Extension to other agricultural tasks and datasets.
- Code: github.com/fmenat/MultiviewCropClassification



A Comparative Assessment of Multi-view Fusion Learning for Crop Classification

IGARSS
19.07.2023

Francisco Mena (1,2), **Diego Arenas** (2), **Marlon Nuske** (2), **Andreas Dengel** (1,2)

1 University of Kaiserslautern-Landau (RPTU)

2 German Research Center for Artificial Intelligence (DFKI)

Paper ID: 2532
email: menatoro@rptu.de



MV Learning

Which is the best multi-view fusion approach for the case study?

- Input, Feature, Decision, Ensemble

