Application of Image Texture Feature Distribution on Agriculture Field Type Classification

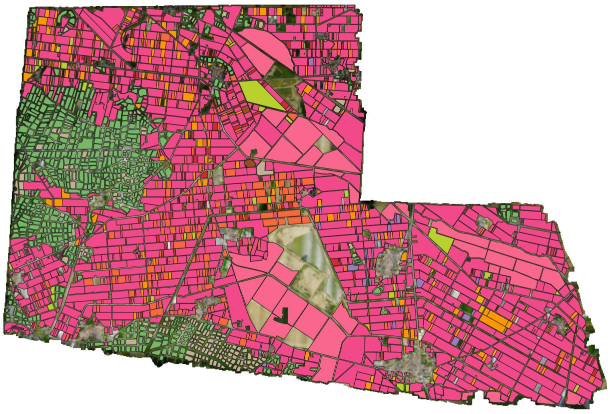
Chun-Han Lee[1], Kuang-Yu Chen[2], **Li-yu Daisy Liu[1*]**

1 Department of Agronomy, National Taiwan University, Taipei 106, Taiwan
2 Environment and Agriculture Division, GEOSAT Aerospace & Technology Inc., Taiwan

Field Classification

- Direct usage:
  - Real-time reports of landuse and landcover (crop maps)
- Indirect usages:
  - Monitoring crop growth/health in large area
  - Prediction for the yield and harvest time
  - Prescription for precision farming
  - Crop insurance …

Images from Unmanned Aerial Vehicle (UAV)



Field Classification

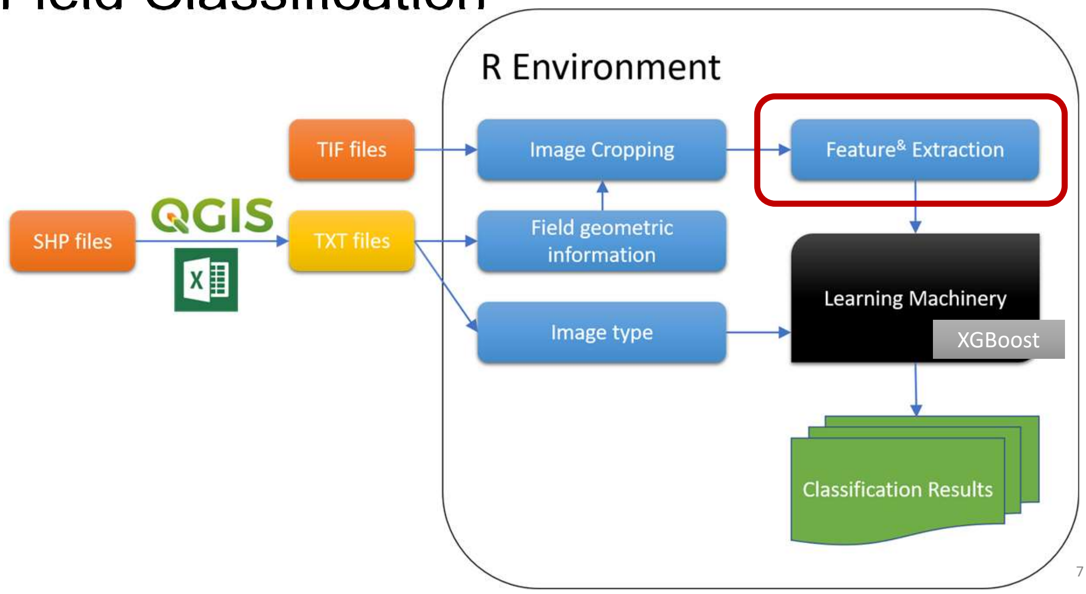
Fruit tree

Greenhouse

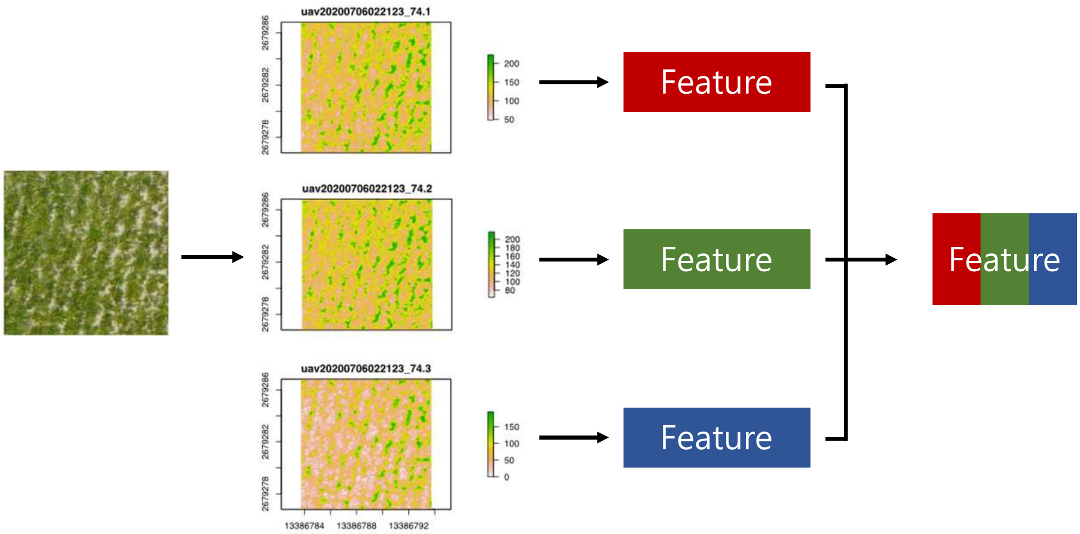
Sugarcane

Maize

Field Classification



Color Feature Extraction



Challenges

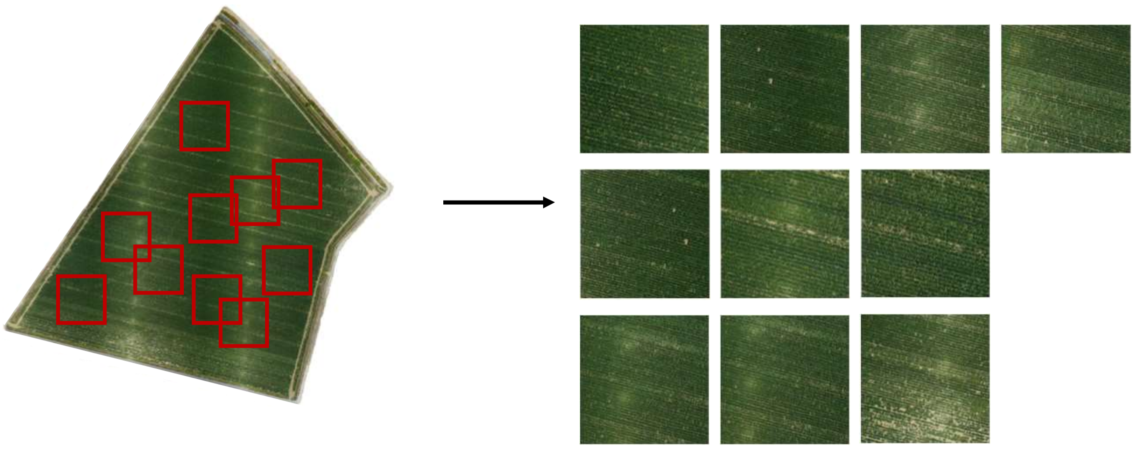
- Highly unbalanced data
- Seasonal effects
- The colors of crops are very similar to each other.

Category	Date			Total
	2019/12	2020/03	2020/06	
Rice	30	533	194	757
Bean	167	214	74	455
Fruit	65	57	72	194
Facility	168	93	118	379
Maize	1980	423	431	2834
Sugarcane	148	41	50	239
Aqua	2512	280	721	3513
Total	5070	1641	1660	8371

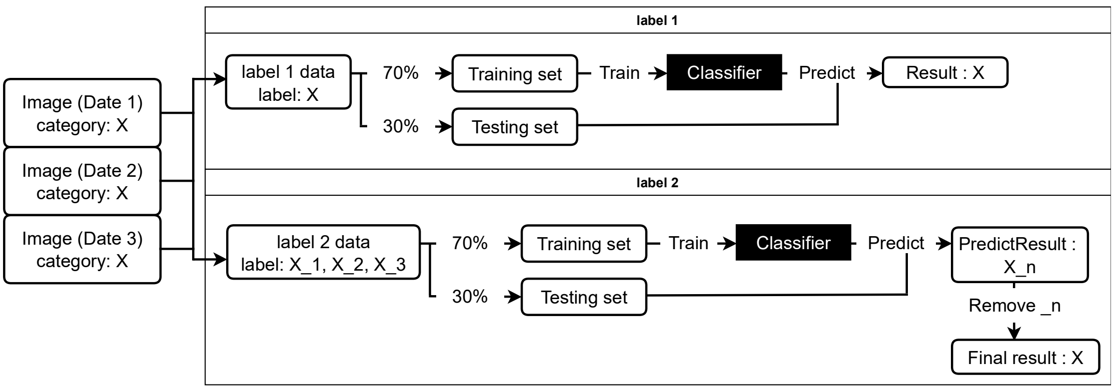
Tricks

- Highly unbalanced data - Take subsamples from the same field
- Seasonal effects - Use time-dependent labeling scheme.
- The colors of crops are very similar to each other - Explore more features that can possibly distinguish the field types.

Unbalance-Data Solution: Resampling



Seasonal-Effect Solution: Labeling



Establish Non-Color Features: Patterns


Fruit tree


Greenhouse


Sugarcane


Maize

"Generally speaking, textures are complex visual patterns composed of entities, or subpatterns, that have characteristic brightness, colour, slope, size, etc. "

Rosenfeld, A. (1976). *Digital picture processing*. Academic press.

Establish Non-Color Features: Patterns

Gray level image

1	4	3	1	5
2	2	5	2	4
1	2	3	1	4
2	1	4	1	1
1	5	5	5	3

Dist = 1

Angle = 0°

GLCM

	1	2	3	4	5
1	1	1	0	3	2
2	1	1	1	1	1
3	2	0	0	0	0
4	1	0	1	0	0
5	0	1	1	0	2

the Gray-Level Co-occurrence Matrix (GLCM)

Traditional Usage of GLCM

GLCM 1

0.5	**0**
0	**0.5**

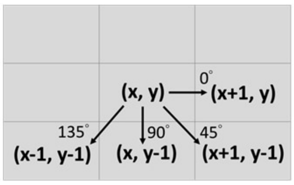
GLCM 2

0.25	**0.25**
0.25	**0.25**

$mean(GLCM\ 1) = 0 \times 0.5 + 1 \times 0.5 = 0.5$

$mean(GLCM\ 2) = 0 \times (0.25 + 0.25) + 1 \times (0.25 + 0.25) = 0.5$

Proposed Usage of GLCM



Angle, θ **(°)**	$cos(\theta)$	$-sin(\theta)$	**Distance,** *dist*
0	1	0	1
45	$1/\sqrt{2}$	$-1/\sqrt{2}$	$\sqrt{2}$
90	0	-1	1
135	$-1/\sqrt{2}$	$-1/\sqrt{2}$	$\sqrt{2}$

The flattened version of the GLCM:

- When presenting a GLCM by its row vectors, $G_{dist}^{\theta} = (\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_K)^T$.
- Then the $K^2 \times 1$ GLCMv is defined as $(\boldsymbol{v}_1^T, \boldsymbol{v}_2^T, \cdots, \boldsymbol{v}_K^T)^T$.

Proposed Usage of GLCM

1	4	3	1	5
2	2	5	2	4
1	2	3	1	4
2	1	4	1	1
1	5	5	5	3

Dist = 1
angle = 90
symm. = F
prob = F

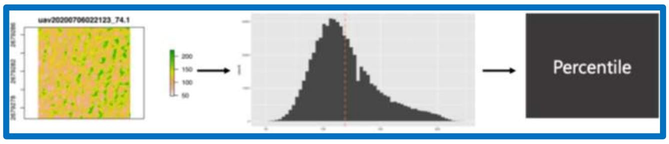
	1	2	3	4	5
1	1	1	0	3	2
2	1	1	1	1	1
3	2	0	0	0	0
4	1	0	1	0	0
5	0	1	1	0	2

Haralick feature:
Mean, Variance

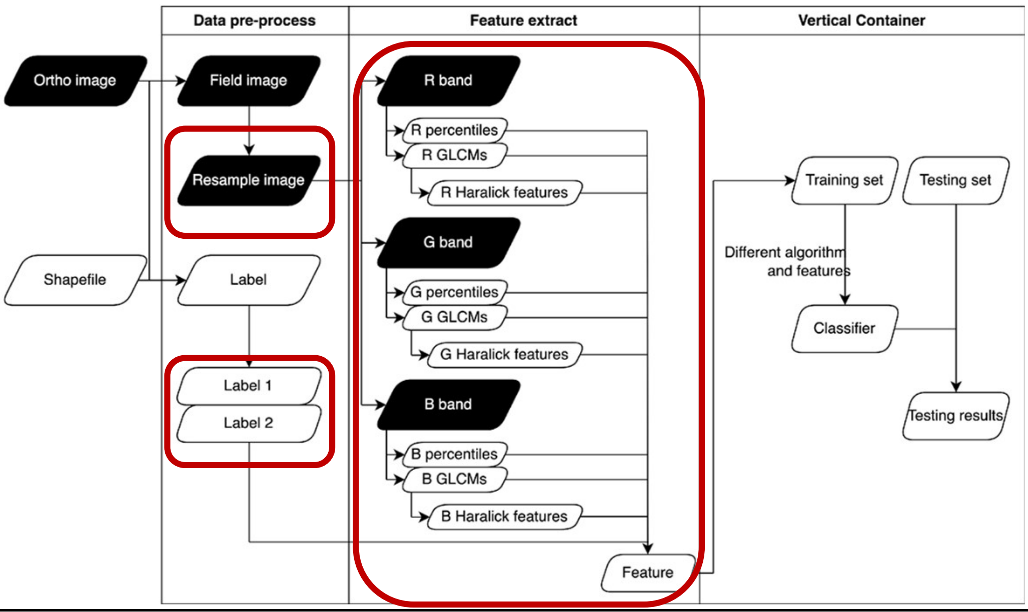
Haralick feature

Flatten GLCMv:
Percentiles

GLCM feature



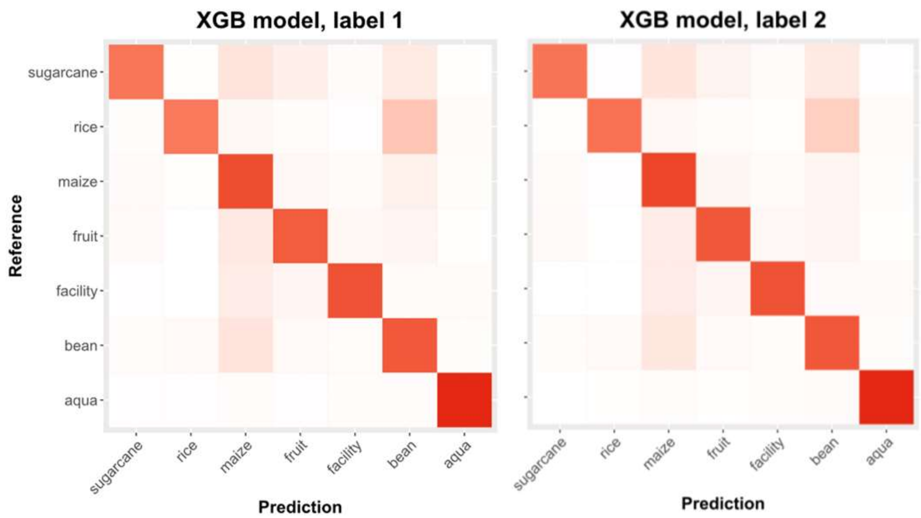
Workflow



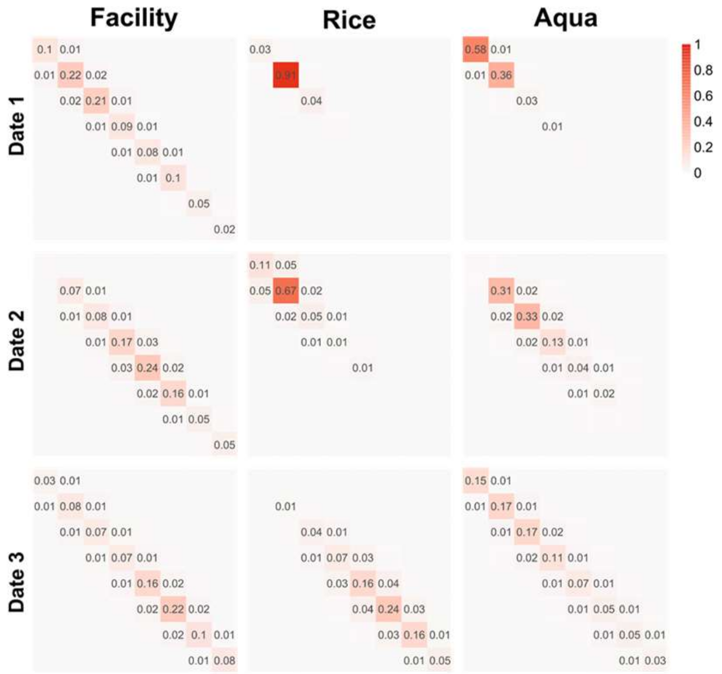
Results - Field Resampling

Category	2019/12		2020/03		2020/06	
	Origin	Resampling	Origin	Resampling	Origin	Resampling
Rice	30	300	533	1066	194	970
Bean	167	1002	214	1070	74	740
Fruit	65	650	57	570	72	720
Facility	168	1008	93	930	118	944
Maize	1980	1980	423	846	431	862
Sugarcane	148	1036	41	410	50	500
Aqua	2512	2512	280	1120	721	721

Performance of Field Classification (Label)



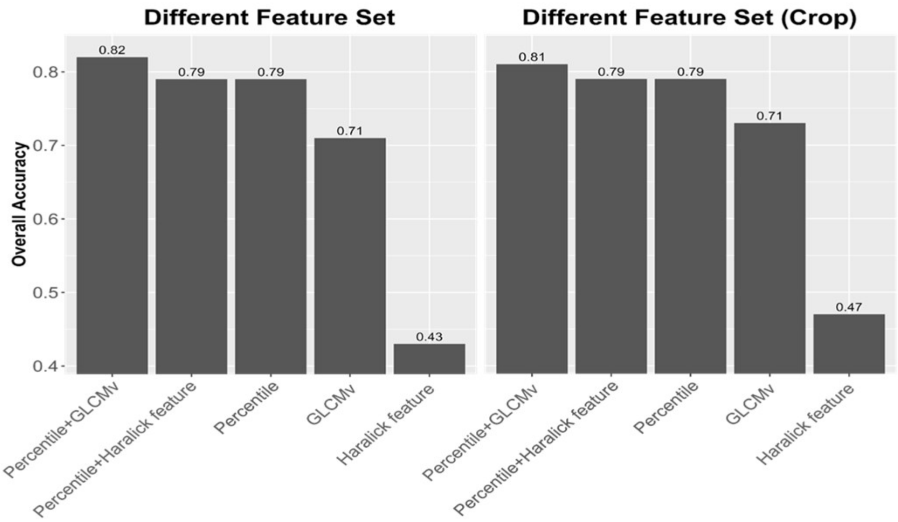
GLCM patterns changed across time



Performance of Field Classification (Label)

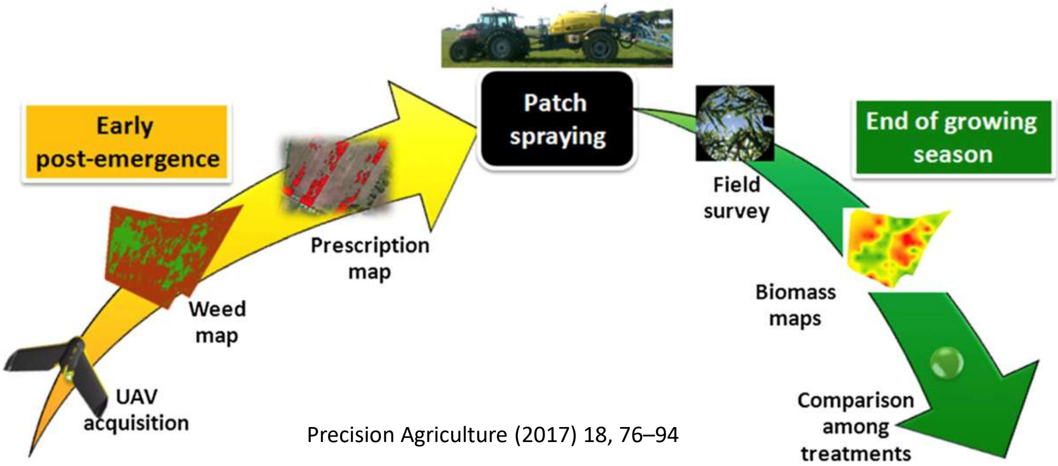
Label	Model	Aqua	Bean	Facility	Fruit	Maize	Rice	Sugarcane
1	CART	0.86	0.71	0.70	0.66	0.71	0.50	0.50
	SVM	0.95	0.80	0.88	0.81	0.83	0.71	0.71
	XGBoost	0.96	0.85	0.90	0.87	0.88	0.81	0.83
2	CART	0.89	0.74	0.77	0.57	0.76	0.54	0.53
	SVM	0.95	0.82	0.88	0.82	0.85	0.73	0.76
	XGBoost	0.97	0.86	0.90	0.89	0.88	0.82	0.84

Performance of Field Classification (Texture)

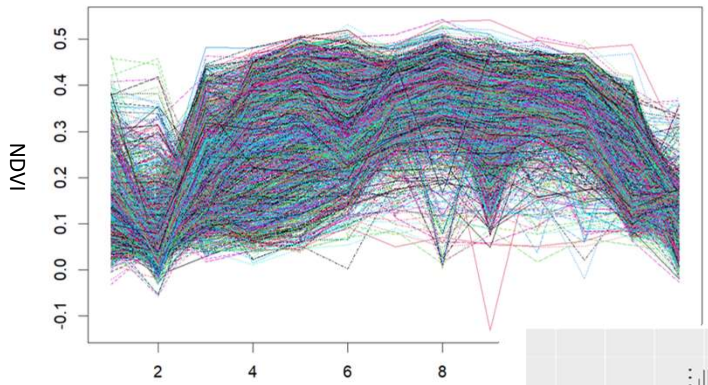


Remarks and Conclusion

Real-time Field Prescription



Precision Agriculture (2017) 18, 76–94

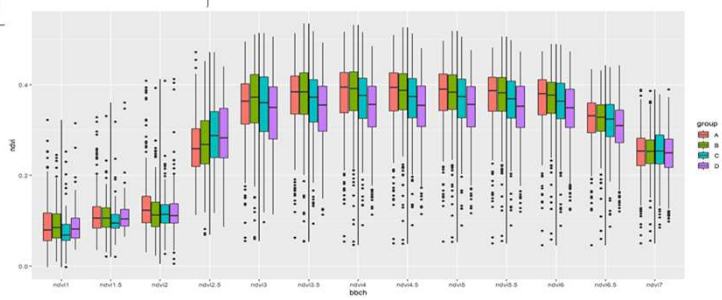


(Unpublished dataset from Chih-Chen Tseng)

By observing the NDVI of paddy rice fields:

- Large noises were perhaps due to time differences of transplanting

- After corrected by the growth stages the NDVI's were more aligned to each other.



Conclusion

- Exploiting the repetitive patterns within the field images inspired us to process through resampling to address the unequal number of samples, and facilitating the use of texture feature.
- The temporal disparities in crop types were crucial in influencing classification outcomes from our observations.
- The XGBoost algorithm, rooted in ensemble learning, outperformed CART or SVM in accuracy and computing times.
- The classification model discerned Taiwan's fragmented and diverse agricultural lands.

Acknowledgement

國立臺灣大學生物資源暨農學院
College of Bioresources and Agriculture, National Taiwan University

先端智農實驗室
Apex Agri-Intelligence Lab

GEOSAT 經緯航太科技股份有限公司
GEOSAT AEROSPACE & TECHNOLOGY INC.