

Microsoft Research

Summit 2022

# Harnessing the Power of AI for Animal Phenotyping



WISCONSIN  
UNIVERSITY OF WISCONSIN - MADISON

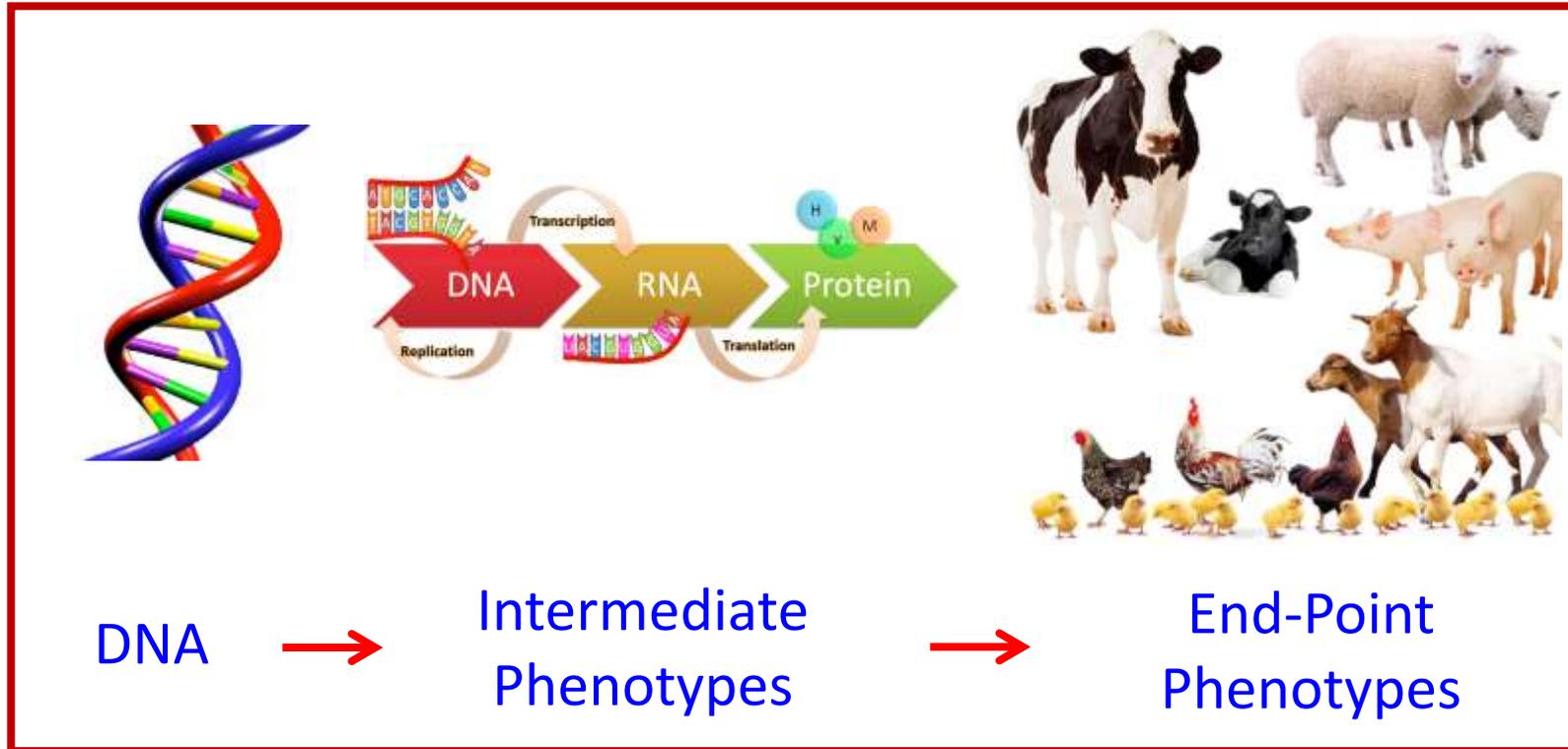
*Joao Dorea*

*Assistant Professor*

*Department of Animal and Dairy Sciences*

*Department of Biological Systems Engineering*

# Omic Technologies

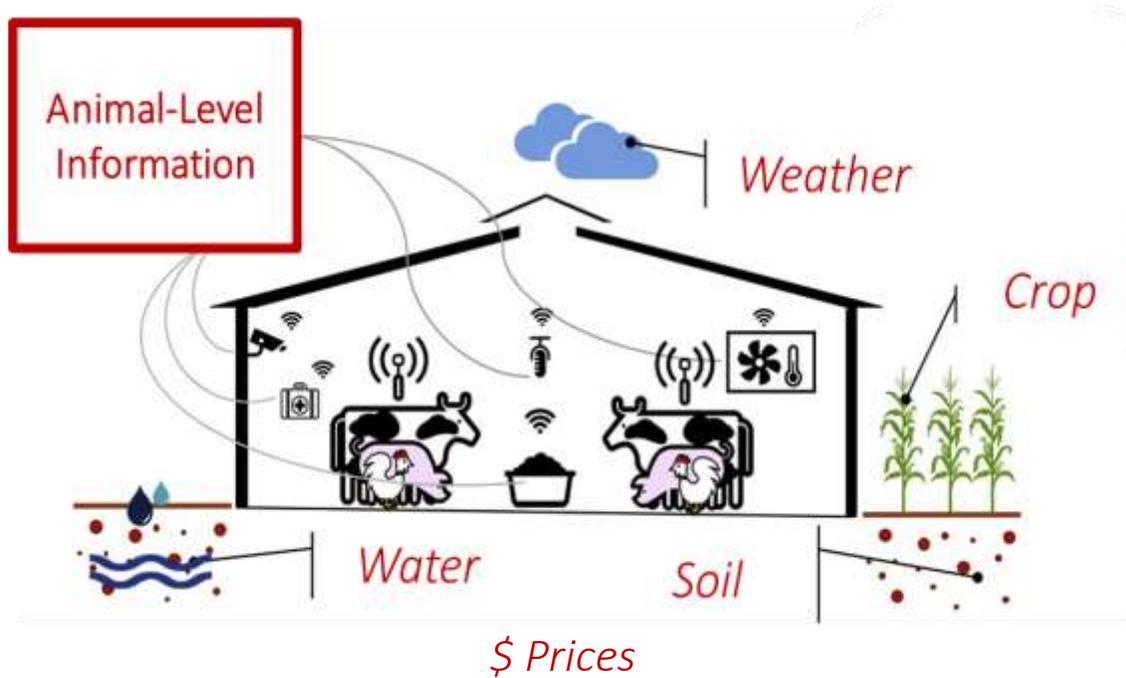


Genomics, Transcriptomics,  
Proteomics, Metabolomics,  
Epigenomics, Microbiomics, etc.

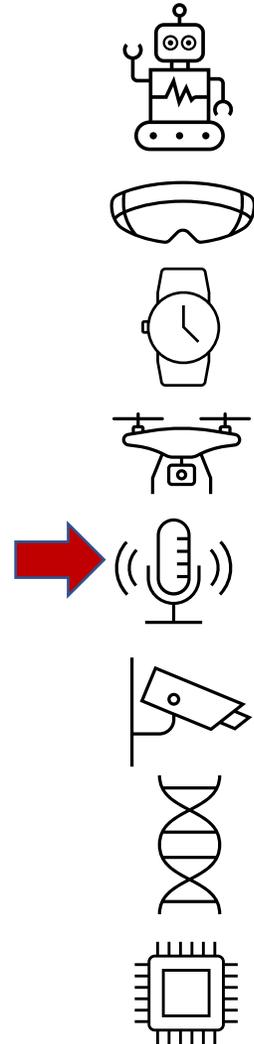
High-Throughput  
Phenotyping (HTP)  
“Phenomics”

# Sensing Technologies to Collect Individual Animal Data

- For large-scale phenotyping:  
“sensing technology is the solution”

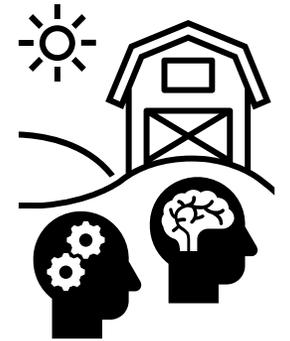
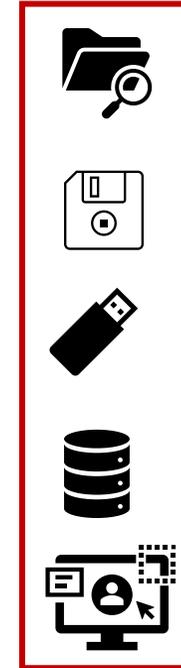


## Multi-Sensor Systems



Animal Behavior  
Traceability  
Infectious Disease  
Body Composition  
Methane Emission  
Social Interaction

## Ag Data

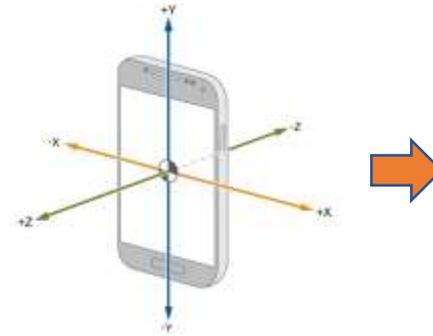
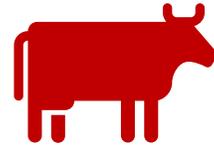
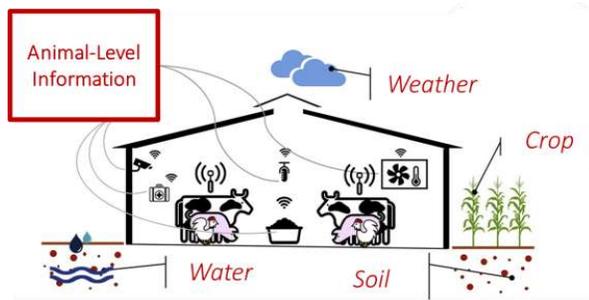


On-farm  
management  
Decisions  
(short-term \$)

Genetic  
Selection  
(long-term \$)

# Why Computer Vision Systems?

Can I use other sensing technologies?



Grazing



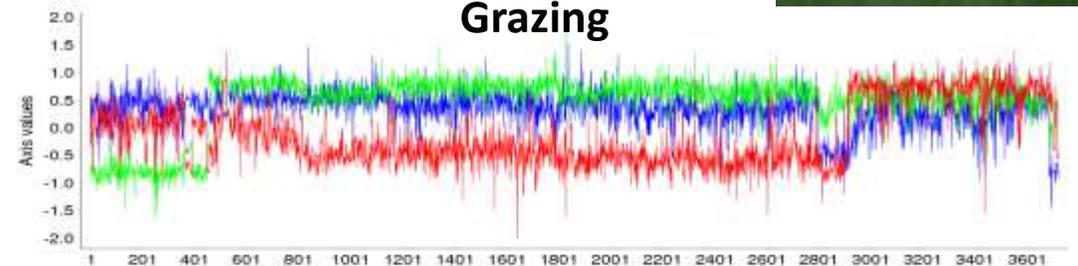
Non-Grazing



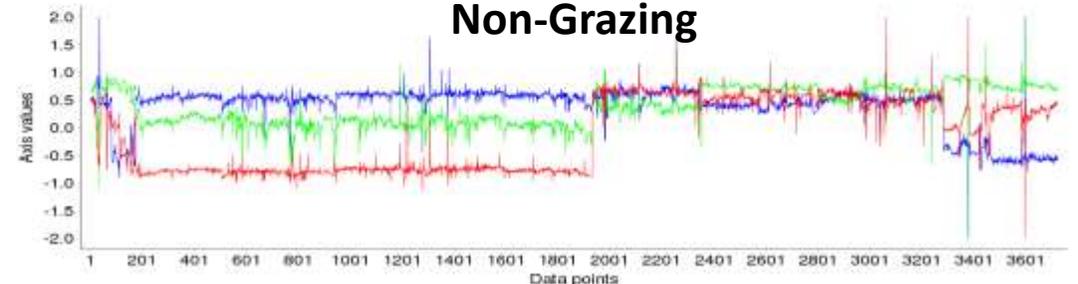
**A single image can be extremely informative!**  
*...it can go beyond your primary interest!*



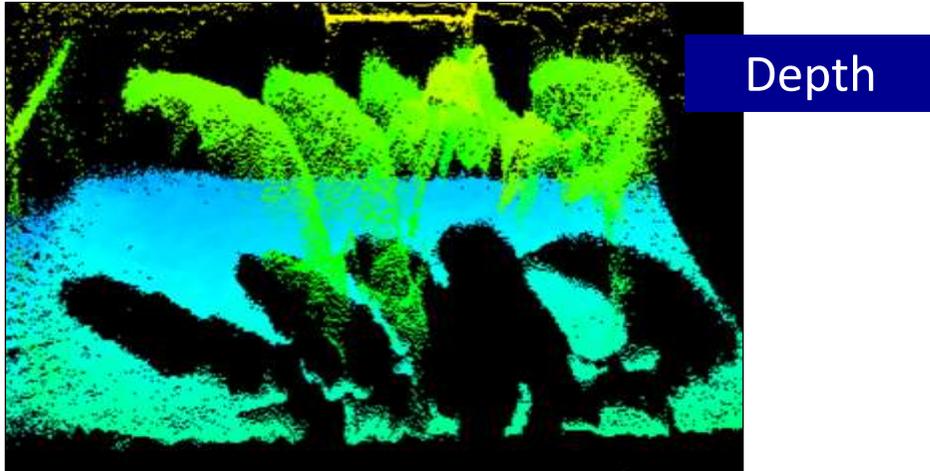
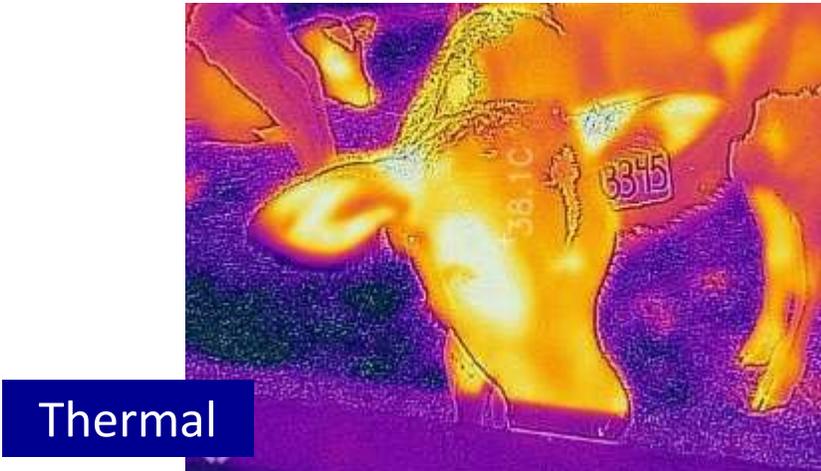
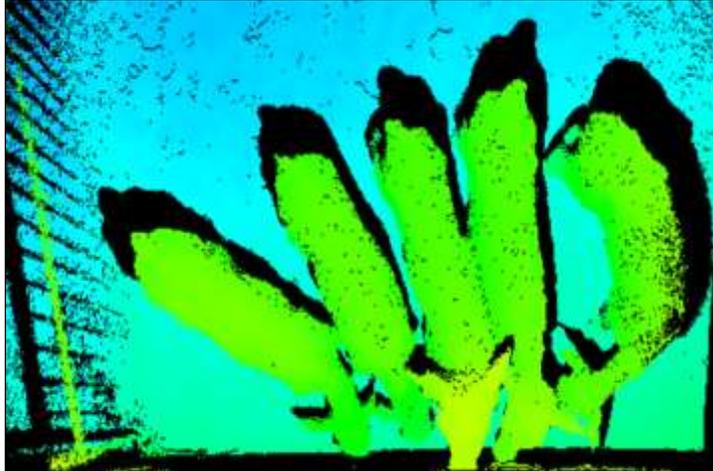
Grazing



Non-Grazing

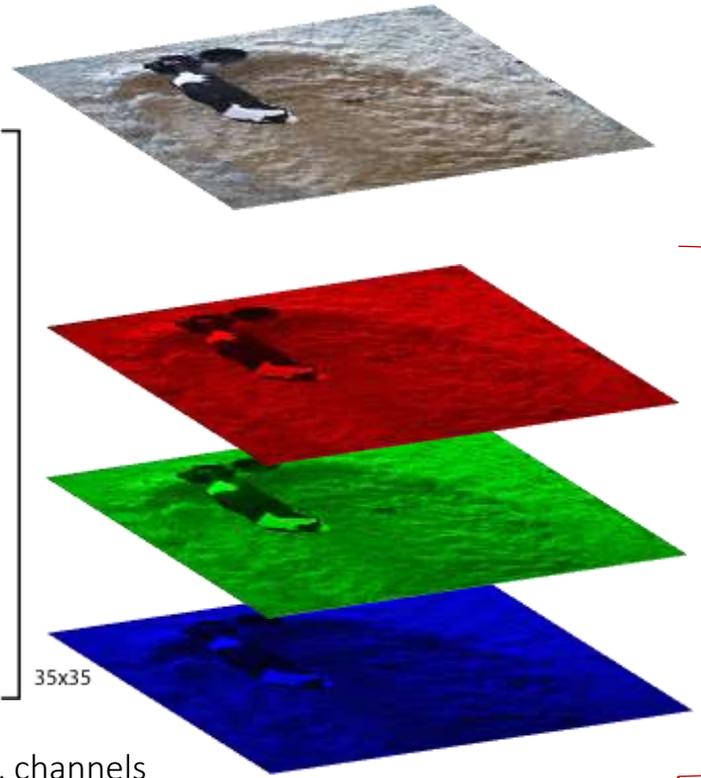
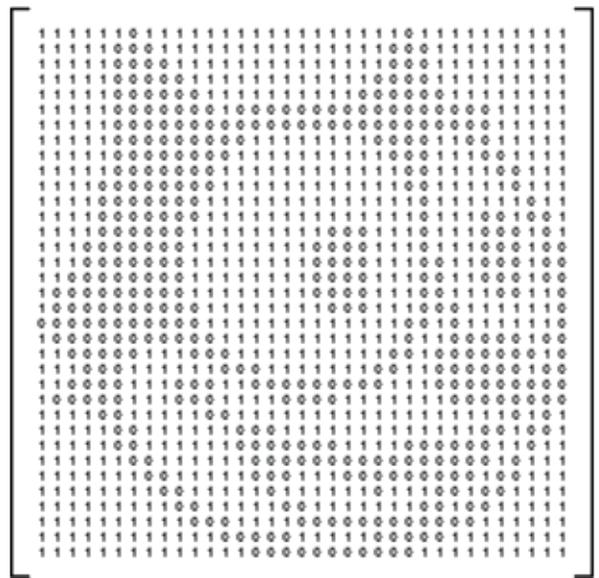
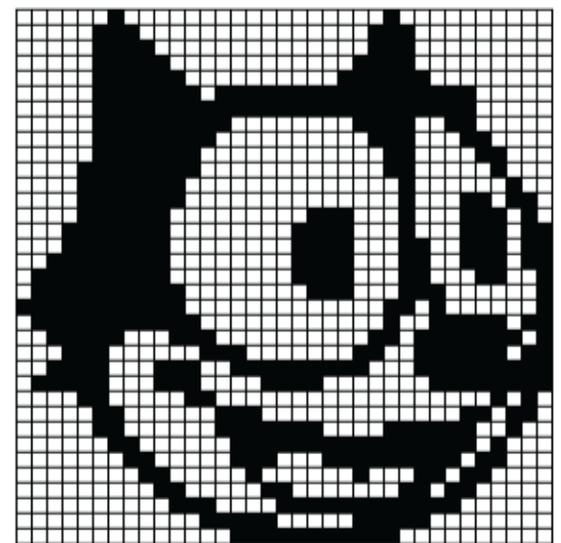


# Computer Vision Systems in Livestock Farming



# Behind the image

- Complex dataset to manage



35 x 35  
Dimension: Number of images, width, height, channels  
(1,35,35,1)

35 x 35  
Dimension: Number of images, width, height, **channels**  
(1,35,35,3)

RGB

# Computer Vision Systems: Image Analyses

- **Complex dataset to analyze**
- **Variety of tasks: Deep Learning Algorithms (CNNs)**

## 2D Image Classification

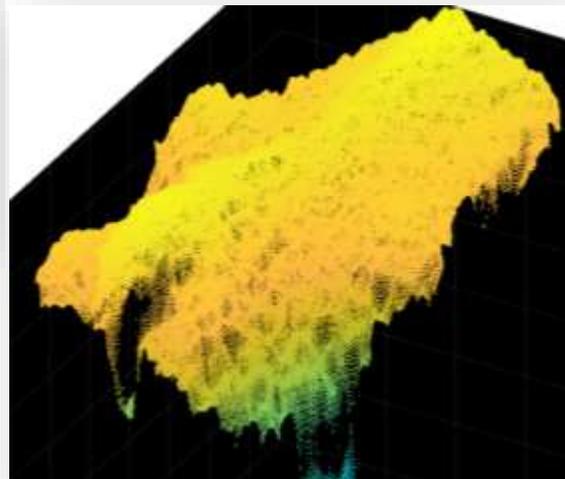
- No calf
- Calf



## Semantic Segmentation



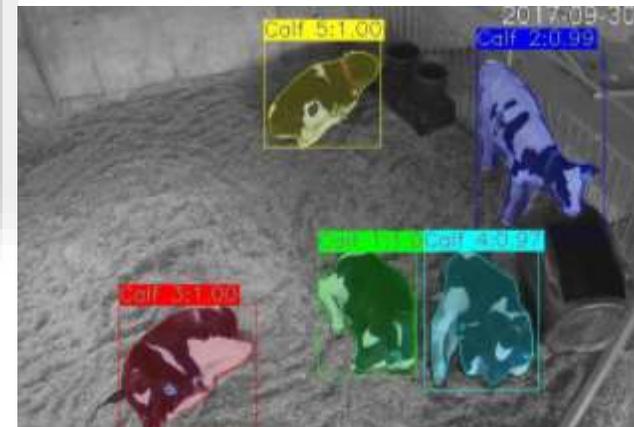
## 3D Image Classification



## Object Detection



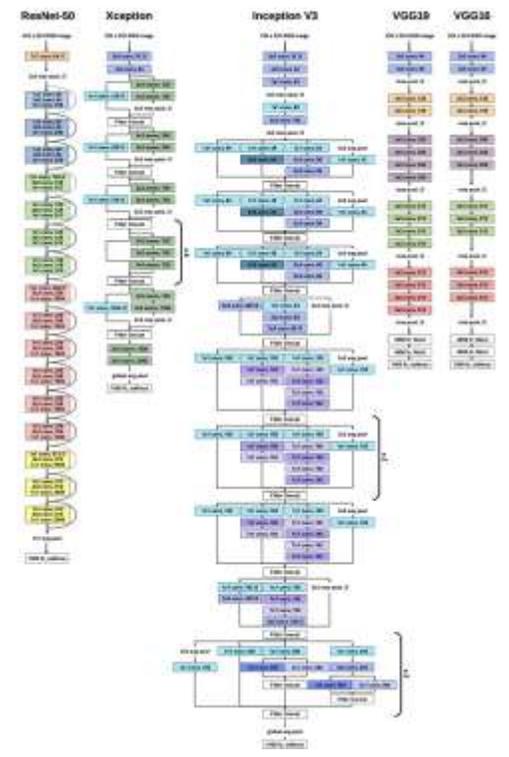
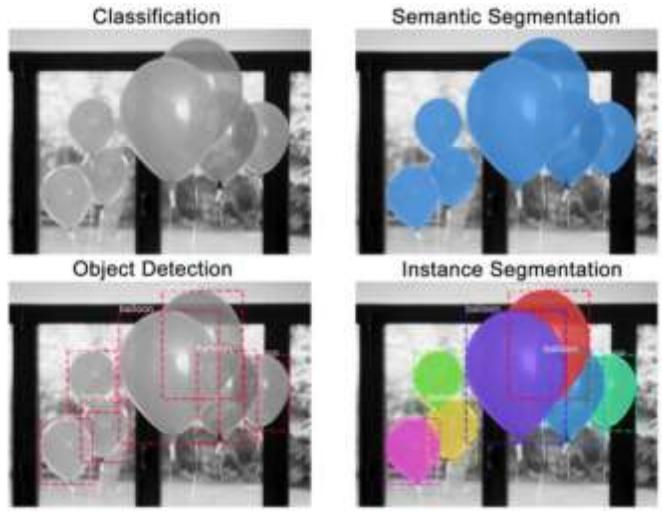
## Instance Segmentation



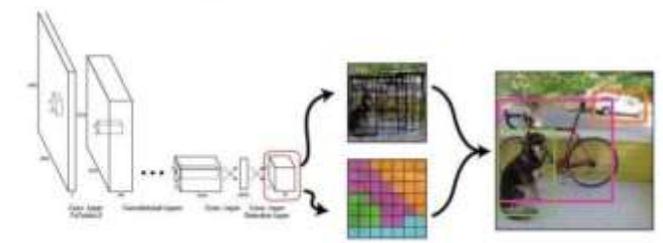
## Image Generation



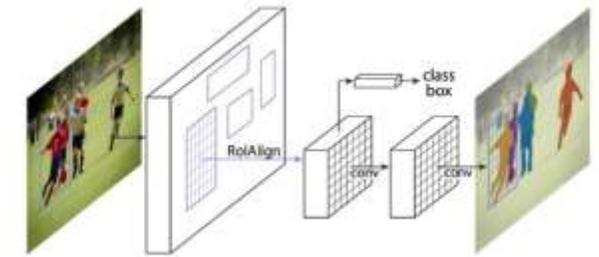
# Computer Vision Systems: Image Analyses



YOLO: You Only Look Once



Object Detection: YOLOv5  
Redmon et al. (2016)



Instance segmentation: Mask-RCNN  
He et al. (2017)

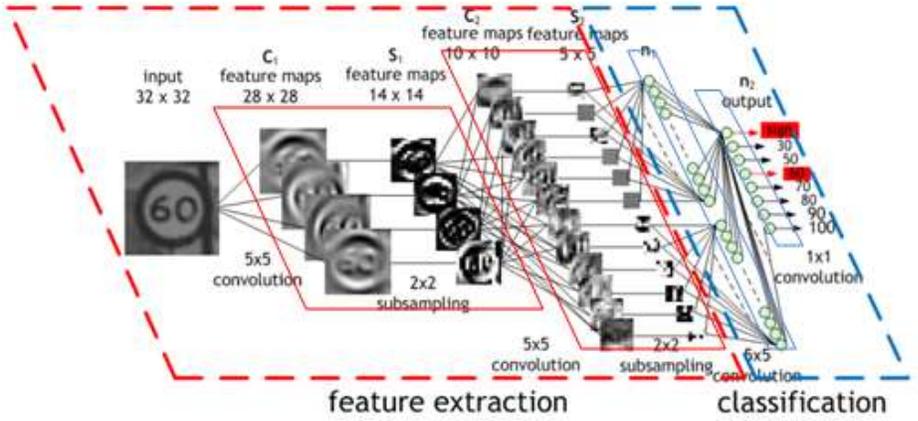
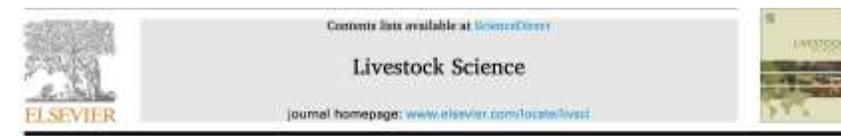


Image Classification:  
VGG16, Resnet, Xception...



A review of deep learning algorithms for computer vision systems in livestock

Dario Augusto Borges Oliveira<sup>a</sup>, Luiz Gustavo Ribeiro Pereira<sup>a,b</sup>, Tiago Bresolin<sup>a</sup>, Rafael Ehrlich Pontes Ferreira<sup>a</sup>, Joao Ricardo Reboucas Dorea<sup>a,b</sup>

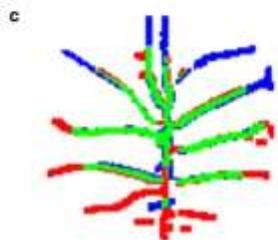
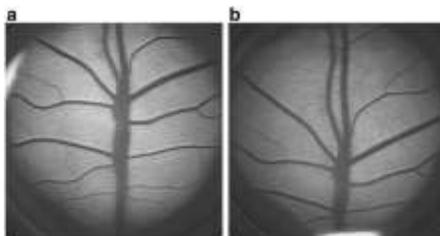
<sup>a</sup> Department of Animal and Dairy Sciences, 1675 Observatory Drive, 286 Animal Sciences Building, Madison, WI 53706-1200  
<sup>b</sup> Instituto Dairy Cattle, Av. Engenheiro de Nascimento, 640 - Aertopos, Jate de Fera - MG, 36038-250, Brazil

# Phenotyping: **Animal Identification**

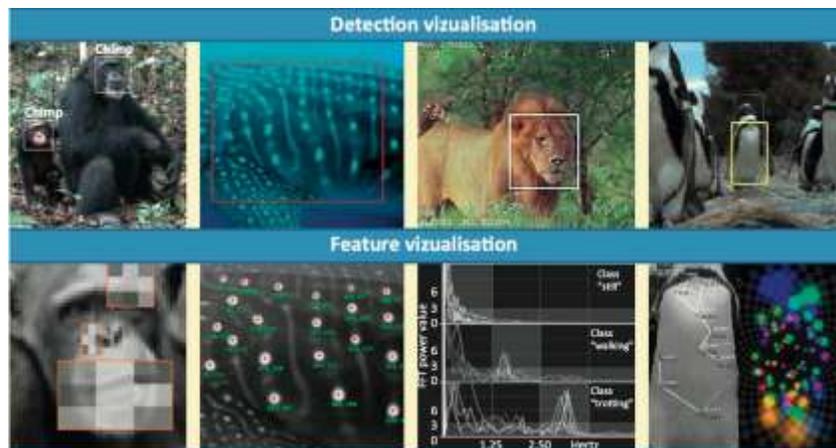
- It is the **first step** for almost all Computer Vision Systems in livestock
- Animal identification and **traceability** is central to ensure food security, transparency, and consumer trust
  - Current animal identification systems rely on ear tags / RFID technology
  - Recently, 2D images are being used as an alternative



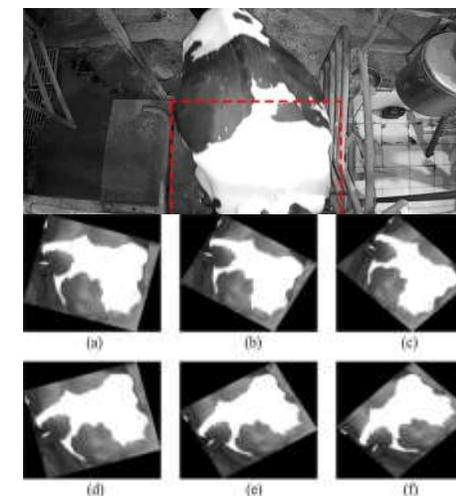
*Petersen et al., 1922  
Journal of Dairy Science*



*Allen et al., 2008  
Livestock Science*

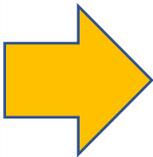
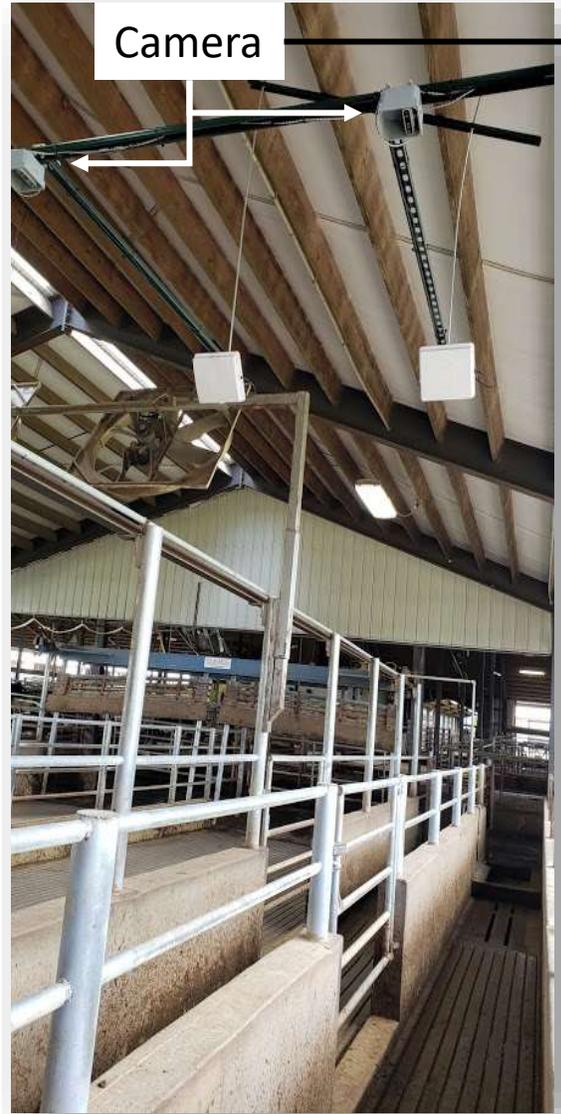


*Kuhl and Burghardt, 2013  
Trends in Ecology & Evolution*



*Li et al., 2017  
Computer and Electronics in Ag.*

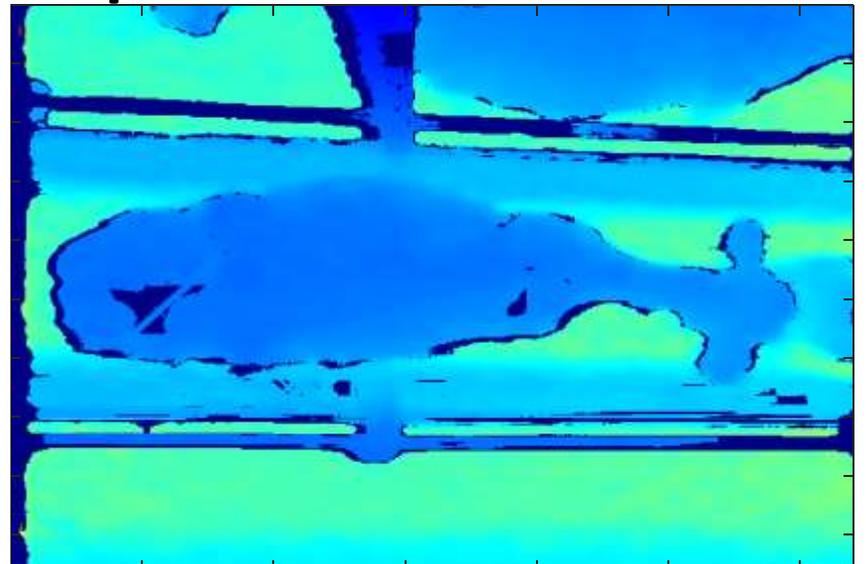
# Large-Scale + Unrestrained Animals + Large unlabeled datasets



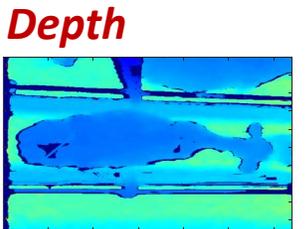
**Infrared**



**Depth**

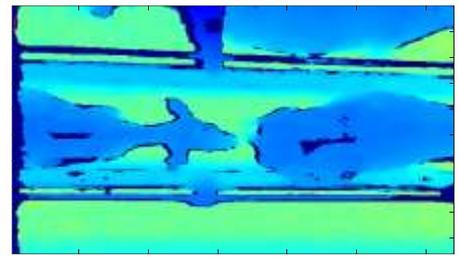


# Automation: Cloud-Computing Framework – Unrestrained Animals

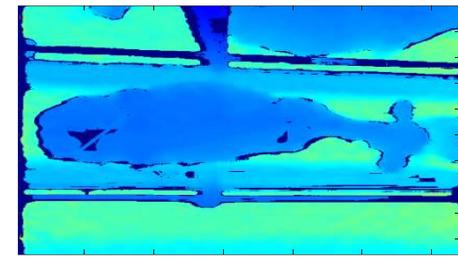


## 1<sup>st</sup> Step: Image Classification

Bad



Good



*Xception (Chollet, 2017)*  
*2D CNN*  
*Accuracy = 96%*  
*52,247 total*       $\longrightarrow$       *19,592 selected*

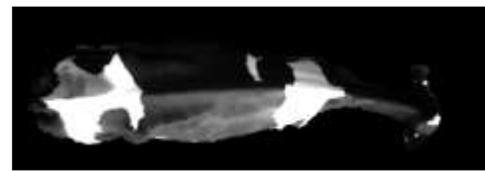
*If good:*

## 2<sup>nd</sup> Step: Image Segmentation (Mask)



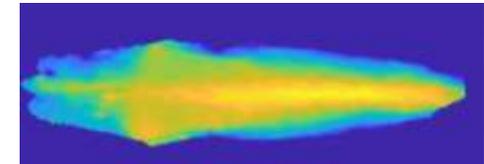
*U-Net (Ronneberger et al., 2015)*  
*2D CNN*  
*Intersection Over Union = 0.93*

## 3<sup>rd</sup> Step: Image Identification (Animal Identification)



*Xception (Chollet, 2017)*  
*2D CNN*

## 4<sup>th</sup> Step: Image Classification (Body Condition Score: 1-5)



*Xception (Chollet, 2017)*  
*2D CNN*

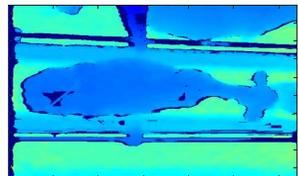
# Automation: Cloud-Computing Framework – Unrestrained Animals



**Infrared**

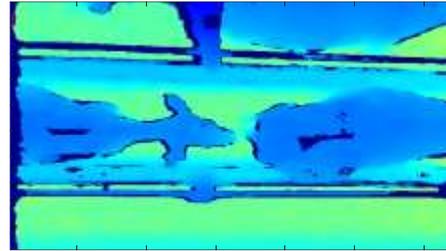


**Depth**

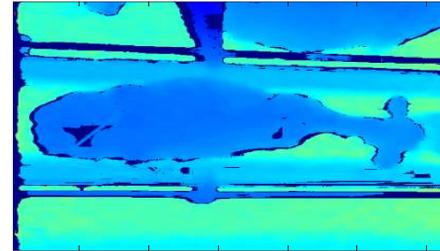


## 1<sup>st</sup> Step: Image Classification

**Bad**



**Good**



*Xception (Chollet, 2017)*

*2D CNN*

*Accuracy = 96%*

*52,247 total*



*19,592 selected*

*If good:*

## 2<sup>nd</sup> Step: Image Segmentation (Mask)

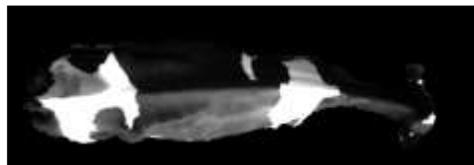


*U-Net (Ronneberger et al., 2015)*

*2D CNN*

*Intersection Over Union = 0.93*

## 3<sup>rd</sup> Step: Image Identification (Animal Identification)



*Xception (Chollet, 2017)*

*2D CNN*

## 4<sup>th</sup> Step: Image Classification (Body Condition Score: 1-5)



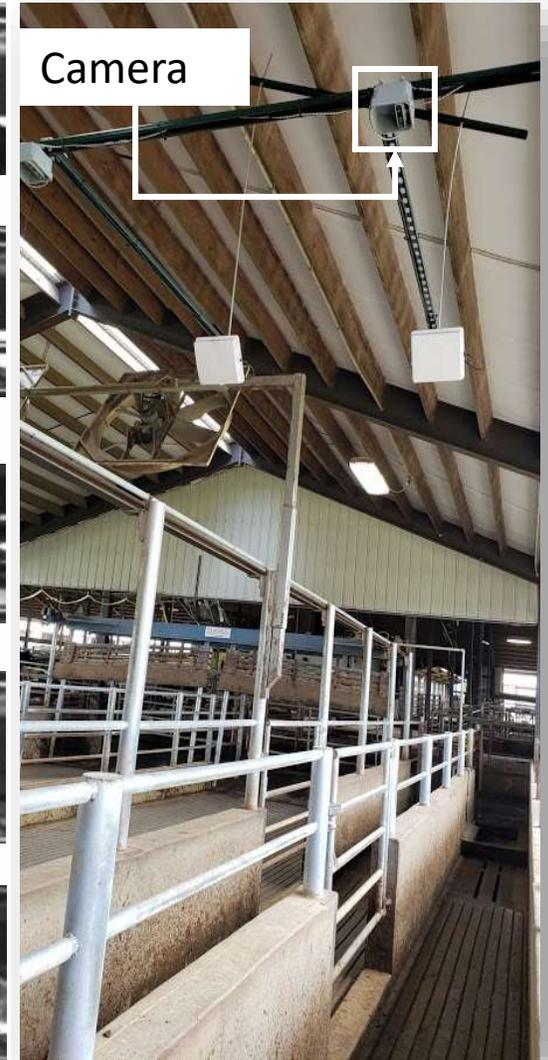
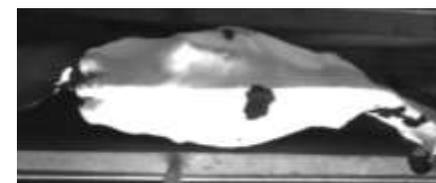
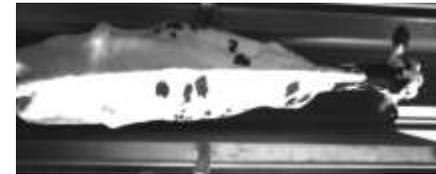
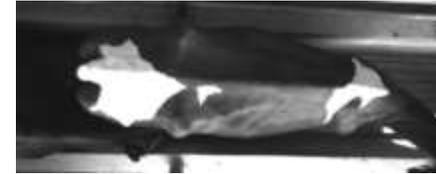
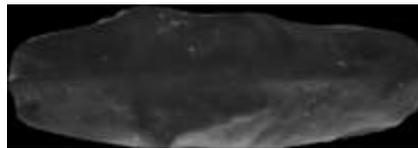
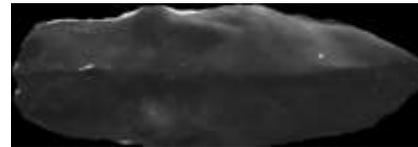
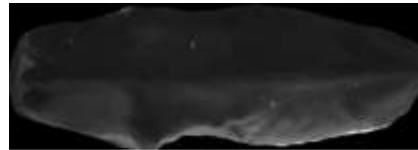
*Xception (Chollet, 2017)*

*2D CNN*

# Animal identification using 2D images

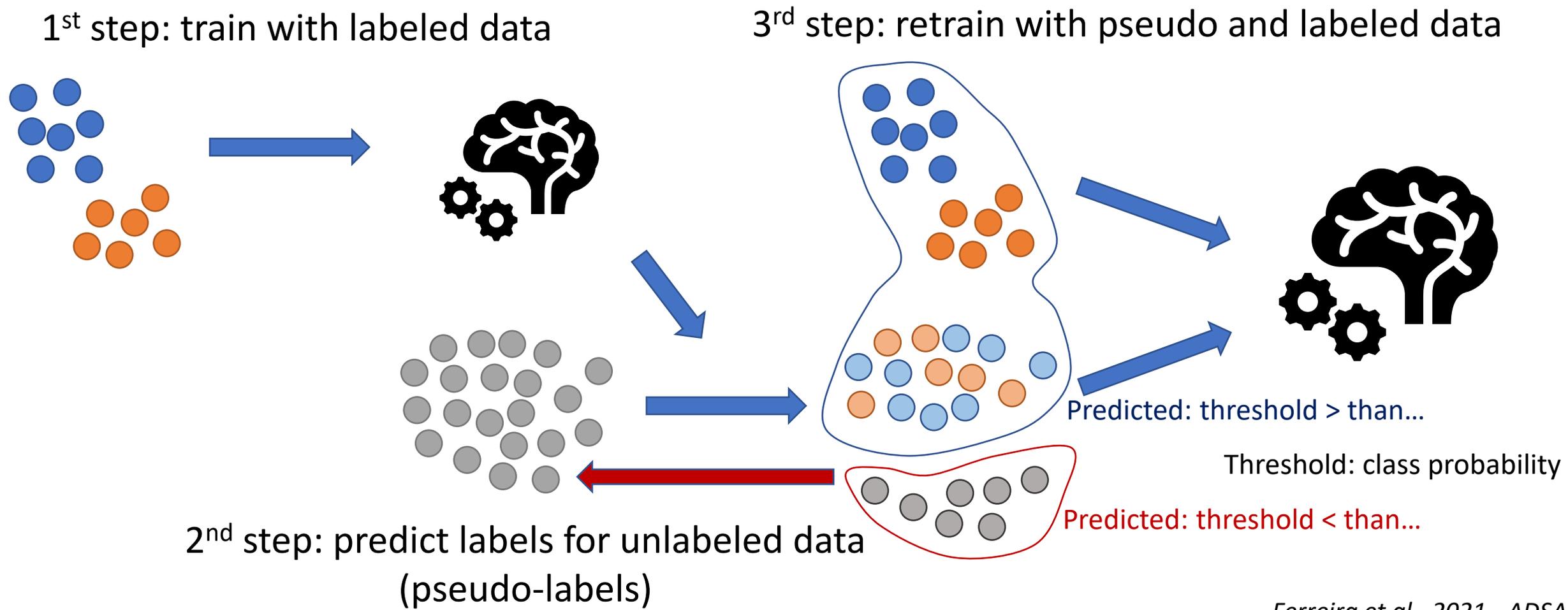
- *92 lactating dairy cows;*
- *Training set: 16,055 images automatically acquired at UW-Madison;*
- *Testing set: 3,680 images test*
- *Deep Learning (CNN; Xception)*
- *Mean Accuracy: 96% to identify individual animals*

High degree of similarity!



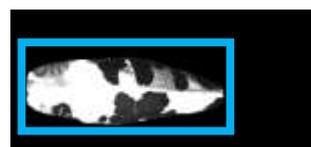
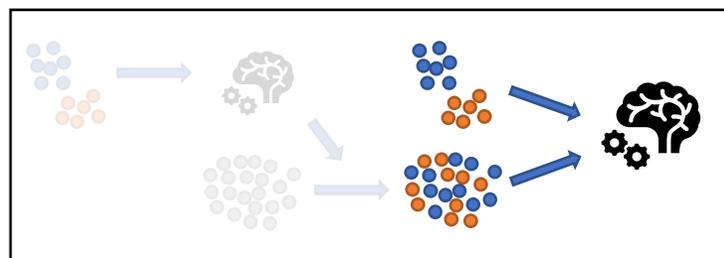
Ferreira et al., 2021 - ADSA

# Semi-Supervised Learning: pseudo-labeling to improve performance of deep neural networks for animal identification



# Semi-Supervised Learning: pseudo-labeling to improve performance of deep neural networks for animal identification

- Used manually labeled and pseudo-labeled images to train a final network



ID 8193



ID 9257



ID 9155

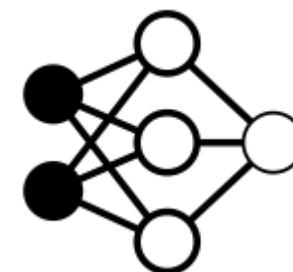
95% > threshold?



ID 8223

88% > threshold?

train



final model

Thresholds: 90%, 95%, 98%, 99% and no threshold

# Semi-Supervised Learning: pseudo-labeling to improve performance of deep neural networks for animal identification



Total dataset: 23,709 -> unlabeled: 20,079 / labeled: 2,354

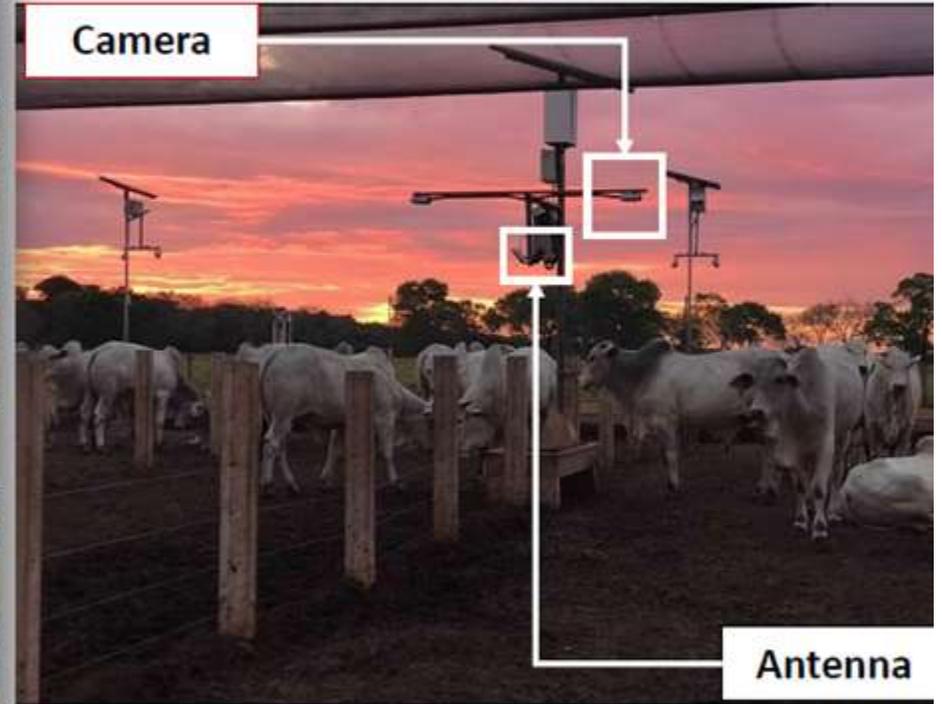
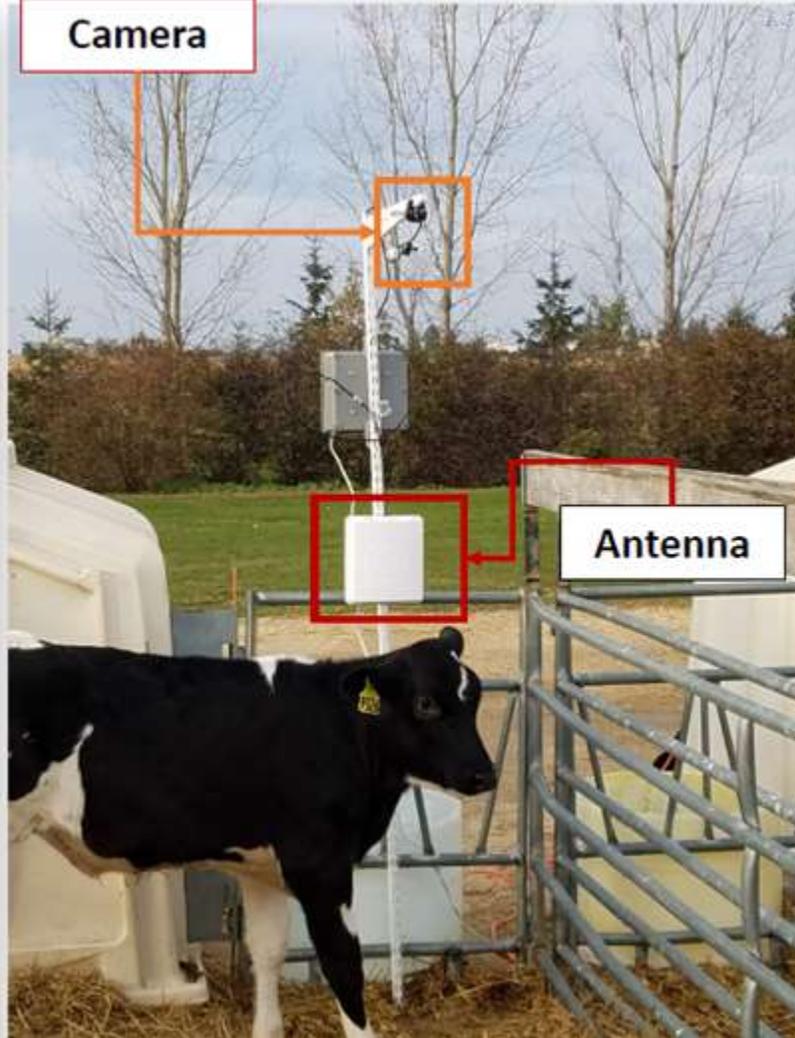
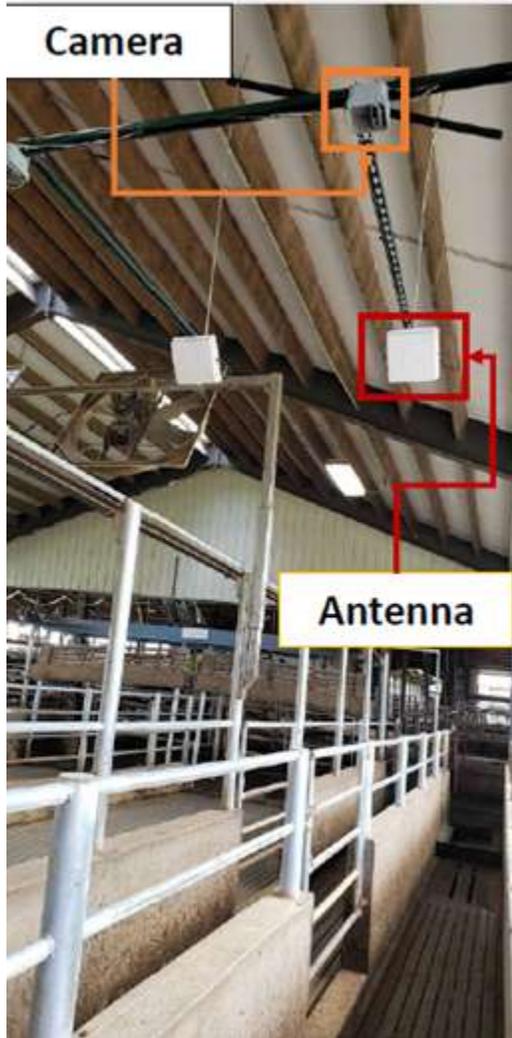
Testing set: 1,180 images

Threshold = 99.999% (defined using validation set with 1,160 images)

Dataset	Baseline training set size	Baseline test accuracy (%)	Final training set size	Final test accuracy <sup>1</sup> (%)	% Test accuracy increase	Initial % of images utilized	Final % of images utilized
	233	33.6	13,822	41.0	22%	1%	68%
	585	51.5	15,476	71.1	38%	3%	74%
	1,177	70.9	20,781	89.2	26%	6%	97%
	1,769	71.4	21,474	89.3	25%	8%	98%
	2,123	74.2	21,990	91.7	24%	10%	99%
	2,354	77.5	21,667	92.0	19%	10%	96%

**increased accuracy from 19% to 38% when compared to the networks trained without pseudo-labeling.**

# Semi-Supervised Learning: RFID can be helpful



# Animal identification

- Is the first step for almost all Computer Vision Systems in livestock:
- Animal identification and traceability is central to ensure food security, transparency, and consumer trust
  - Current animal identification systems rely on ear tags / RFID technology
    - *Labor-intensive, prone to human error or fraud, and can generate stress*
  - Recently, 2D images are being used as an alternative

***It will not work for similar color patterns***



# Animal identification

- 3D animal biometry can be a powerful tool to identify animals with similar color patterns



**To investigate the use of 3D animal biometry as a potential tool for animal identification**



**To investigate how changes in body surface affect animal identification over time**



Using dorsal surface for individual identification of dairy calves through 3D deep learning algorithms

Rafael E.P. Ferreira <sup>a</sup>, Tiago Bresolin <sup>a</sup>, Guilherme J.M. Rosa <sup>a,b</sup>, João R.R. Dórea <sup>a,c,\*</sup>



# Animal identification using 3D images

## 3D images:

**Voxels** (VoxNet; Maturana and Scherer, 2015)

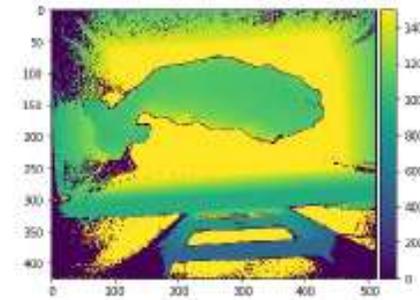
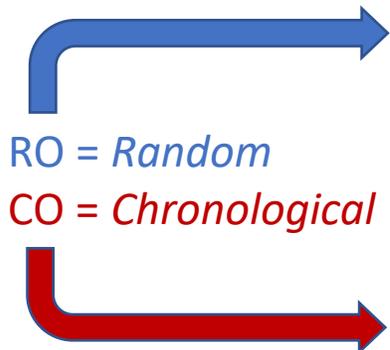
**Point cloud** (PointNet; Qi et al., 2016)

## 2D images:

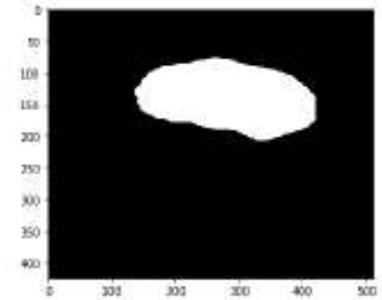
**Depth images**

(VGG16, Xception, Inception v3)

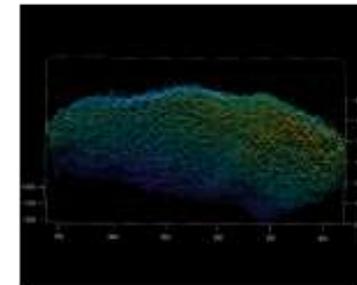
Train-test split	Data representation	Architecture	$F_1$ score
RO <sup>1</sup>	DI <sup>3</sup>	VGG16	0.888
RO <sup>1</sup>	DI <sup>3</sup>	Inception v3	0.904
RO <sup>1</sup>	DI <sup>3</sup>	<b>Xception</b>	<b>0.959</b>
RO <sup>1</sup>	PC <sup>4</sup>	PointNet	0.669
RO <sup>1</sup>	OG <sup>5</sup>	VoxNet	0.880
CO <sup>2</sup>	DI <sup>3</sup>	VGG16	0.718
CO <sup>2</sup>	DI <sup>3</sup>	Inception v3	0.750
CO <sup>2</sup>	DI <sup>3</sup>	<b>Xception</b>	<b>0.804</b>
CO <sup>2</sup>	PC <sup>4</sup>	PointNet	0.429
CO <sup>2</sup>	OG <sup>5</sup>	VoxNet	0.656



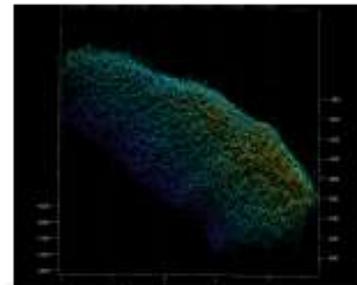
(a) Original depth frame



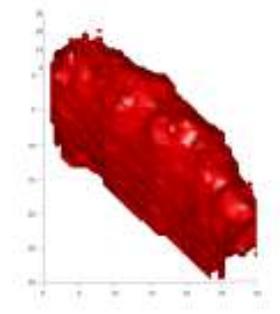
(b) Output from Mask R-CNN



(c) Generated point cloud



(d) Augmented point cloud



(e) Generated occupancy grid

Ferreira et al., 2022

# Animal identification using 3D images

## 3D images:

**Voxels** (VoxNet; Maturana and Scherer, 2015)

**Point cloud** (PointNet; Qi et al., 2016)

## 2D images:

**Depth images**

(VGG16, Xception, Inception v3)

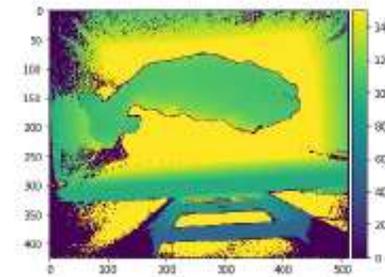
Group	Weeks in training set	Test week	Test set size
No skipping	1 and 2	3	319
	2 and 3	4	254
	3 and 4	5	250
	4 and 5	6	403
Skipping one week	1 and 2	4	254
	2 and 3	5	250
	3 and 4	6	403
Skipping two weeks	1 and 2	5	250
	2 and 3	6	403
Skipping three weeks	1 and 2	6	403



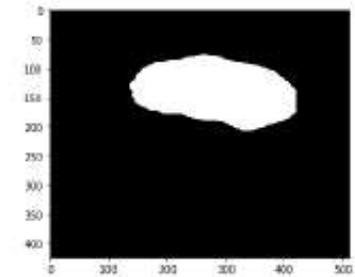
## F1-score

Time interval	Xception	PointNet	VoxNet
No skipping	<b>0.917</b>	0.533	0.917
1 week	<b>0.846</b>	0.551	0.831
2 weeks	<b>0.835</b>	0.441	0.806
3 weeks	<b>0.856</b>	0.282	0.792

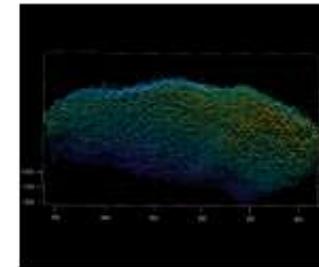
## How frequent should I retrain the algorithms?



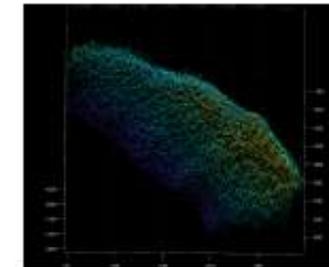
(a) Original depth frame



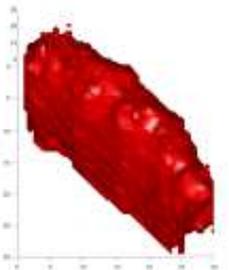
(b) Output from Mask R-CNN



(c) Generated point cloud



(d) Augmented point cloud

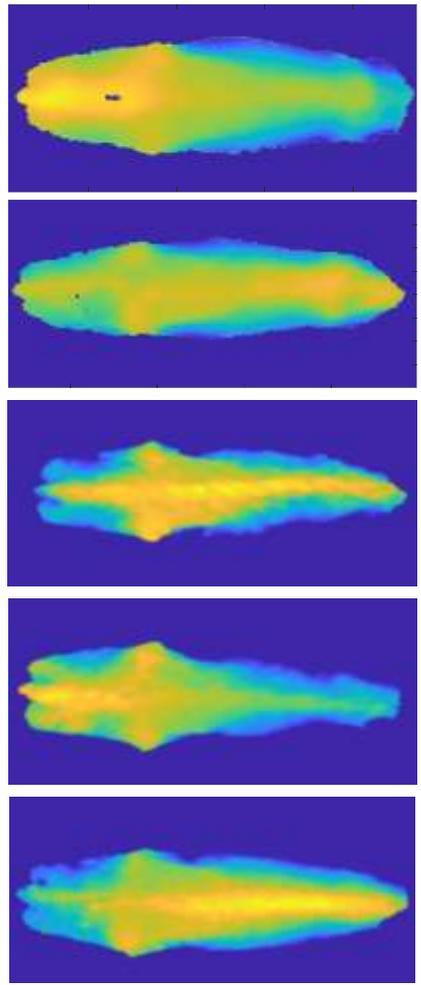
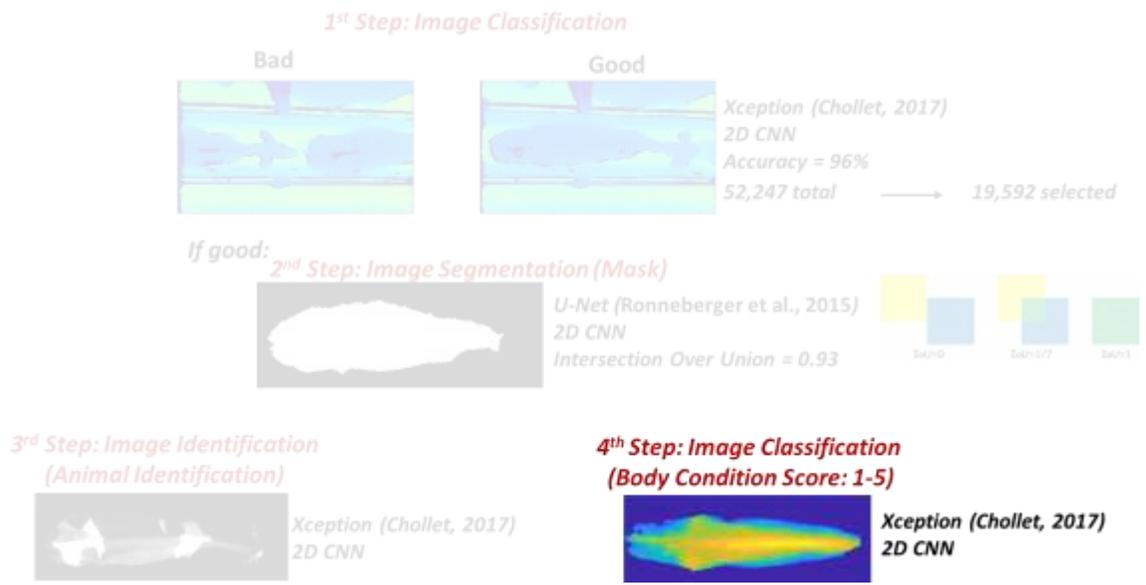


(e) Generated occupancy grid

*\*training set size remained constant*

# Body Condition Score using 3D images

- 59 lactating dairy cows
- Train: 11,943 images
- Test: 651 images
- Deep Learning (CNN; Xception)
- Avg Accuracy: **71%** to classify BCS
- Mean Absolute Error: **0.20**



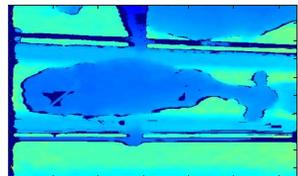
# Automation: Cloud-Computing Framework



**Infrared**

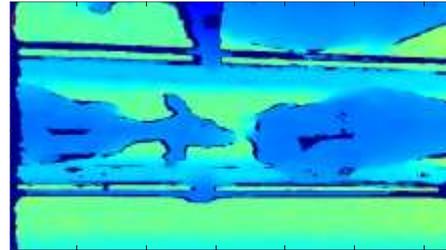


**Depth**

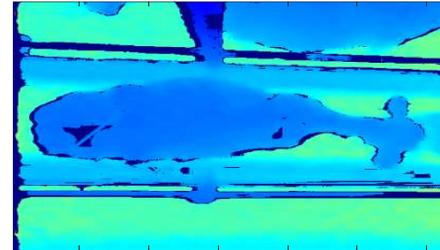


## 1<sup>st</sup> Step: Image Classification

**Bad**



**Good**



*Xception (Chollet, 2017)*

*2D CNN*

*Accuracy = 96%*

*52,247 total*



*19,592 selected*

*If good:*

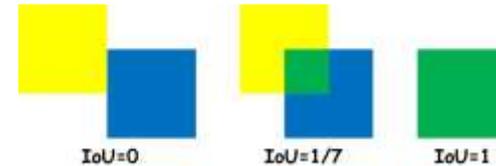
## 2<sup>nd</sup> Step: Image Segmentation (Mask)



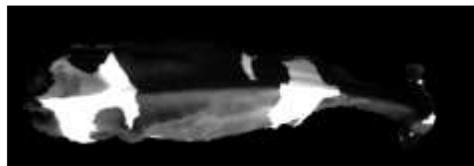
*U-Net (Ronneberger et al., 2015)*

*2D CNN*

*Intersection Over Union = 0.93*



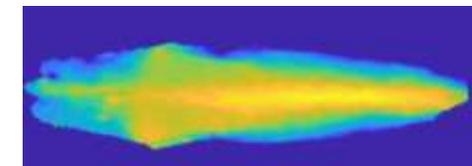
## 3<sup>rd</sup> Step: Image Identification (Animal Identification)



*Xception (Chollet, 2017)*

*2D CNN*

## 4<sup>th</sup> Step: Image Classification (Body Condition Score: 1-5)



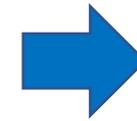
*Xception (Chollet, 2017)*

*2D CNN*

# Early detection of subclinical ketosis in dairy cows

- *Goal: Use prepartum 3D images to predict subclinical ketosis (1-14 DIM)*
- 21, 14 and 7 days prior to calving;
- 76 Holstein cows were individually collected;  
(40 SCK and 36 non-SCK)
- 27,300 top-down 3D images;

**Subclinical  
ketosis**



**Prediction**



**1-14 days**



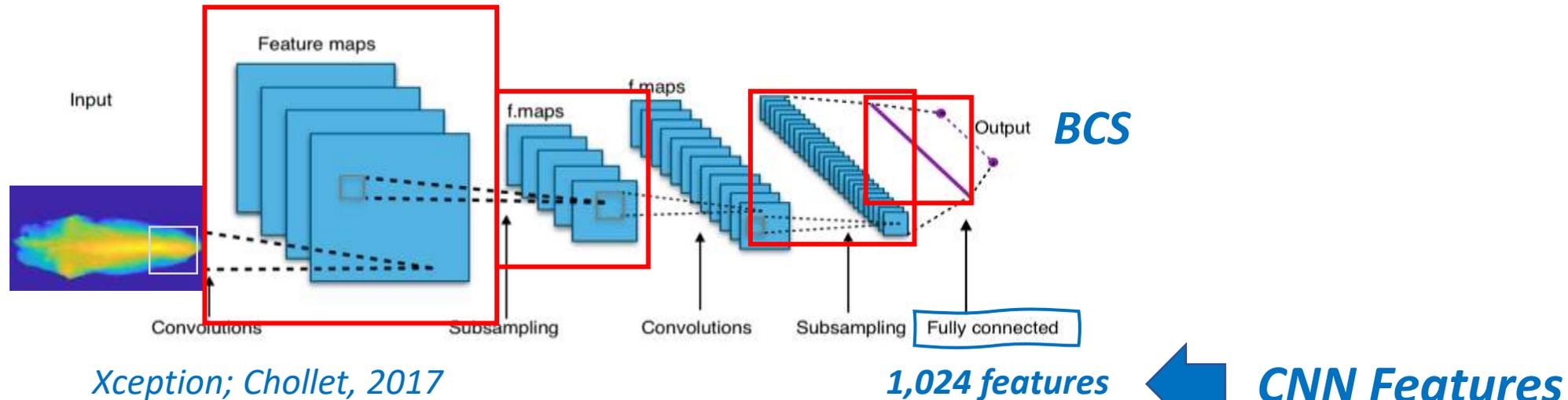
**Postpartum**

**Prepartum**

# Early detection of subclinical ketosis in dairy cows

For each image:

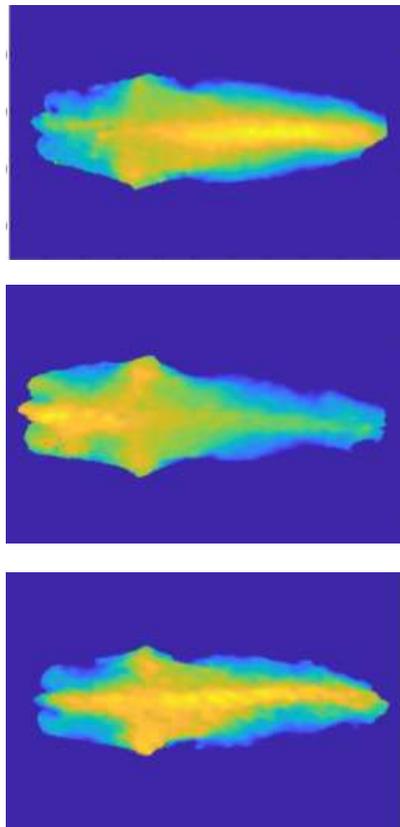
- Biological features (mask size, surface area, volume)
- CNN features (*Xception* architecture, trained to evaluate BCS)



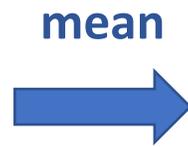
**Biological + CNN Features: total of 1027 features/image;**

# Early detection of subclinical ketosis in dairy cows

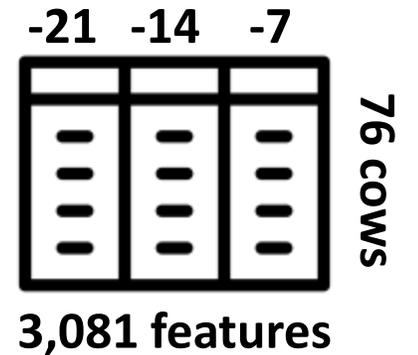
- **Biological** + **CNN** Features: total of 1027 features/image;



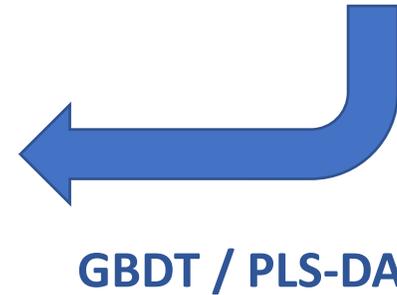
100 images per  
cow per day



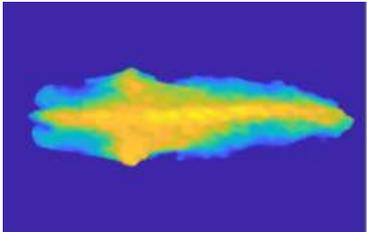
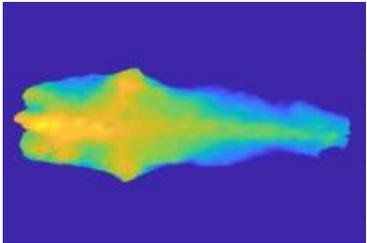
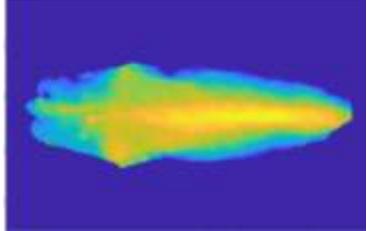
1,027 features per  
cow per day



Subclinical ketosis?  
yes/no



# Early detection of subclinical ketosis in dairy cows



- Predictive Approaches:
  - Gradient Boosting Decision Tree (GBDT)
  - Partial Least-Squares Discriminant Analysis (PLS-DA)
- The dataset was randomly split into training (n=65) and testing set (n=11), and this procedure was repeated 10 times
- Hyperparameters were selected using a 5-fold cross-validation (training set) to maximize F1-score
- Comparison between using our 3,081 features and using 3 BCS evaluations (baseline)

# Early detection of subclinical ketosis in dairy cows

- PLS-DA achieved a mean **precision of 0.65**, **recall of 0.91**, and **F1-score of 0.75**:

Features	Algorithm	Precision (mean $\pm$ stdev)	Recall (mean $\pm$ stdev)	F1-Score (mean $\pm$ stdev)
BCS only	GBDT	0.503 $\pm$ 0.160	0.828 $\pm$ 0.205	0.611 $\pm$ 0.150
BCS only	PLS-DA	0.534 $\pm$ 0.148	0.963 $\pm$ 0.078	0.678 $\pm$ 0.125
Our features	GBDT	0.630 $\pm$ 0.094	0.908 $\pm$ 0.106	0.739 $\pm$ 0.086
<b>Our features</b>	<b>PLS-DA</b>	<b>0.650 <math>\pm</math> 0.090</b>	<b>0.912 <math>\pm</math> 0.102</b>	<b>0.754 <math>\pm</math> 0.081</b>



**+11%**



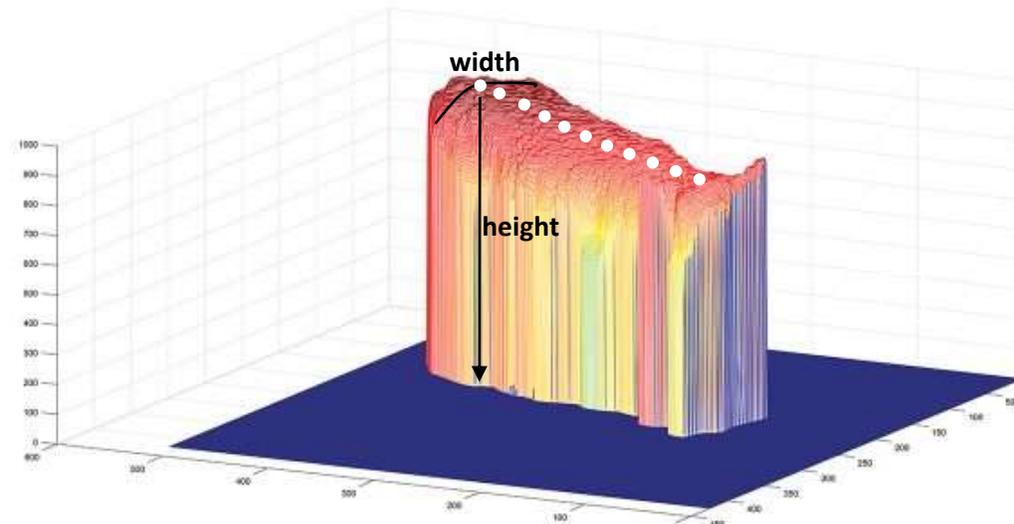
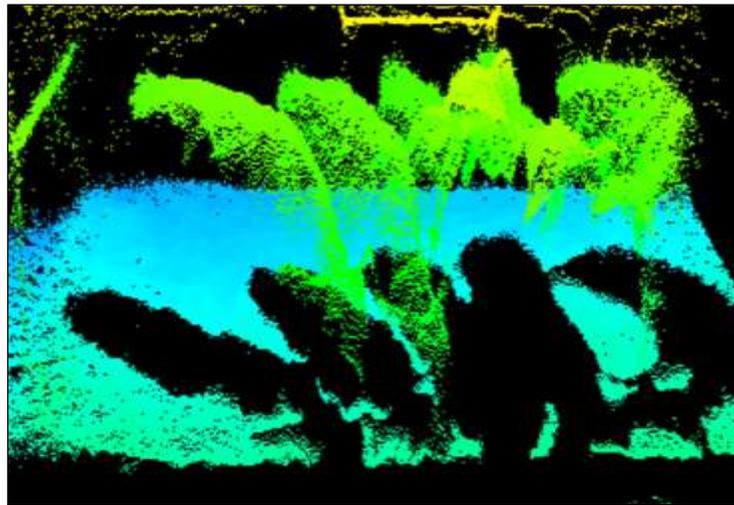
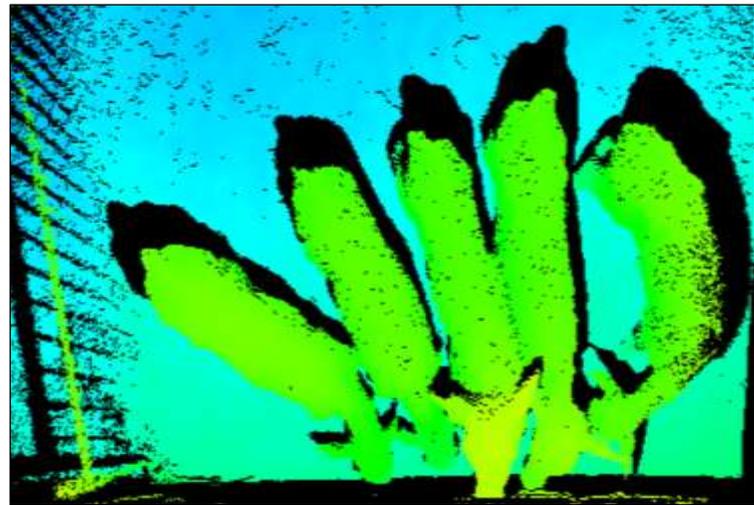
~65% of detected cows  
actually got sick



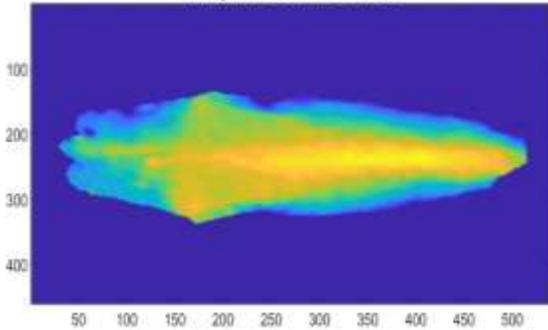
~91% of sick cows  
were early detected

**False positives are better than false negatives!**

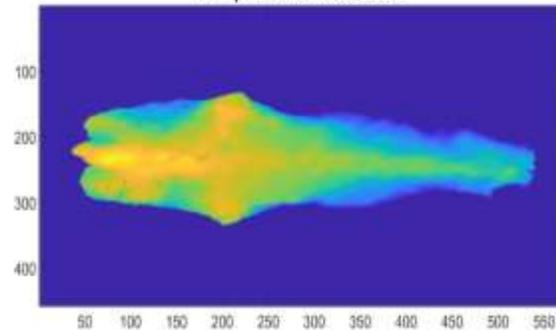
# Monitoring Growth and Body Shape Using 3D Images



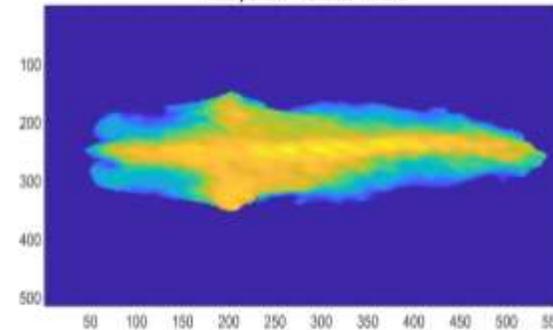
Time point 2 - volume = 114.3 L



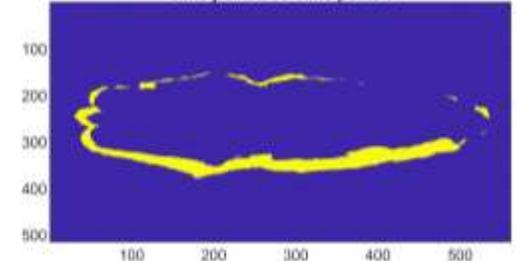
Time point 3 - volume = 91.3 L



Time point 4 - volume = 77.6 L

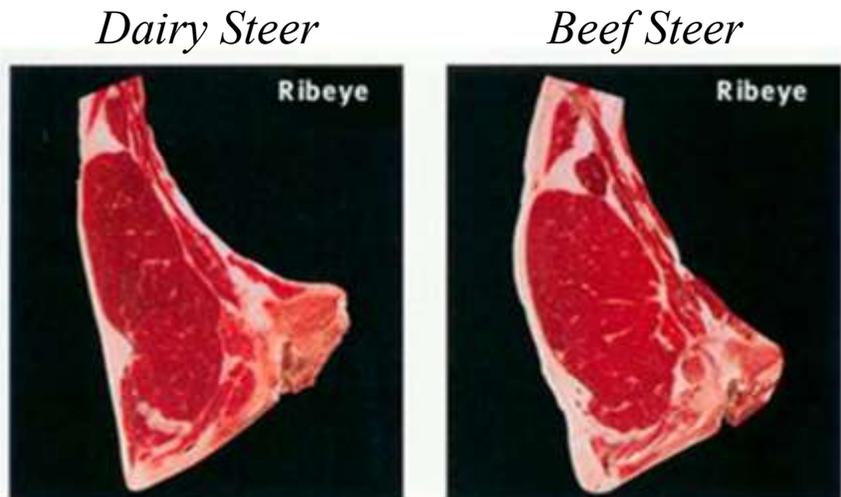


Time point 2 vs Time point 4

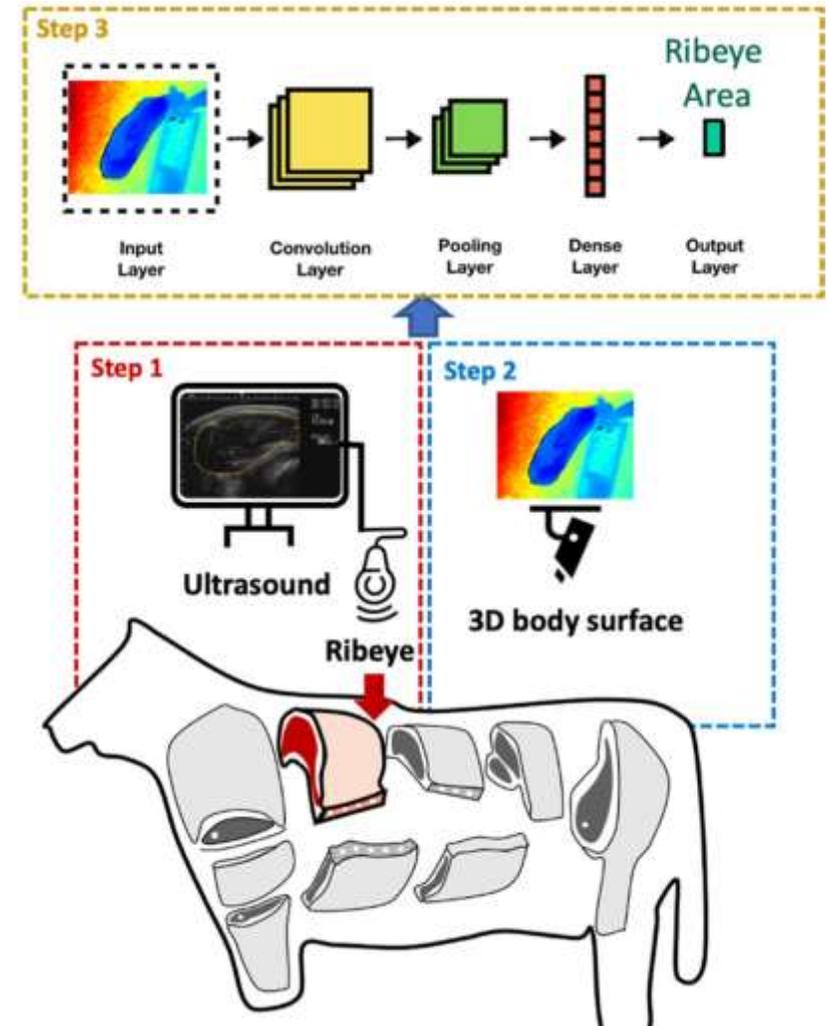


# 3D Body Shape: Predicting Ribeye Area and Shape of Live Calves

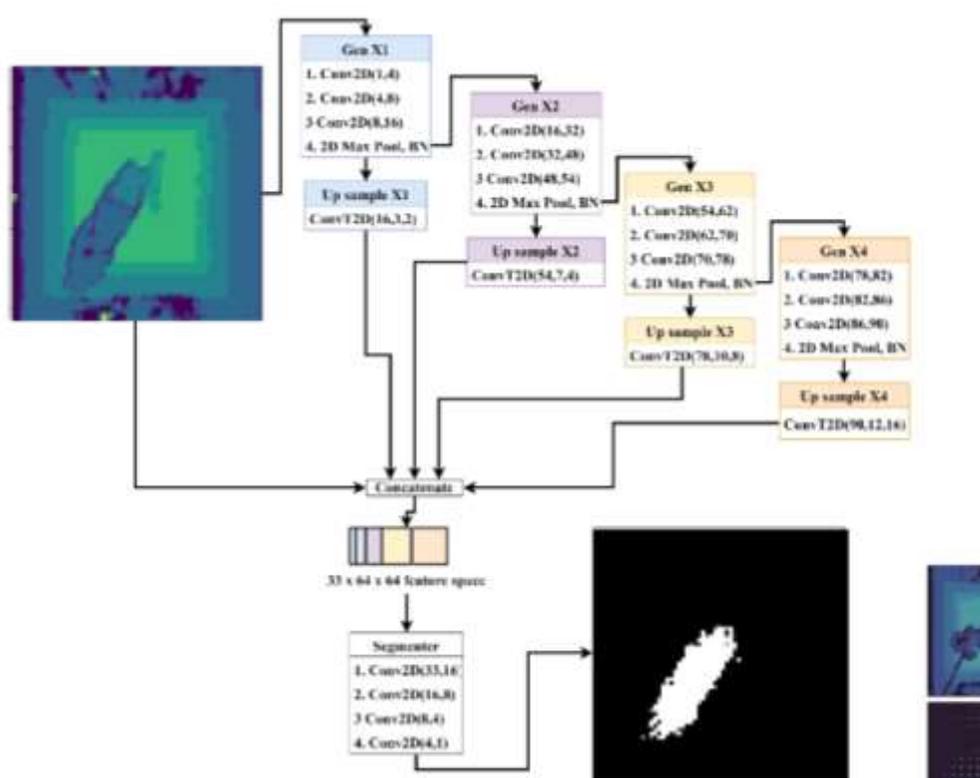
- Beef semen has been used in genetically inferior cows to produce crossbred (**beef x dairy**) animals
- Very important source of income for dairy farmers
- Frequently reported as lacking quality and shape uniformity



Source: Angus Beef Bulleting



# Predicting Ribeye Area and Shape of Live Calves

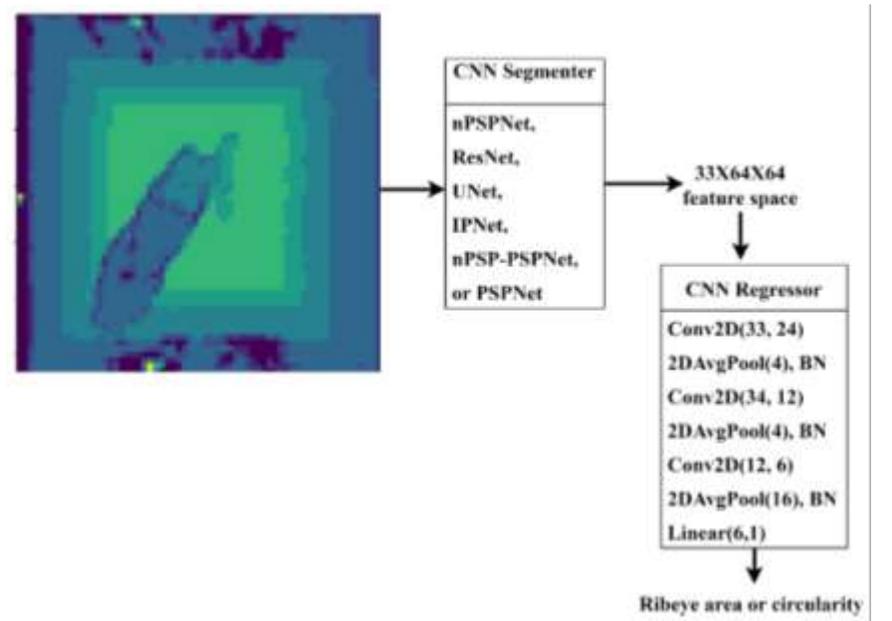
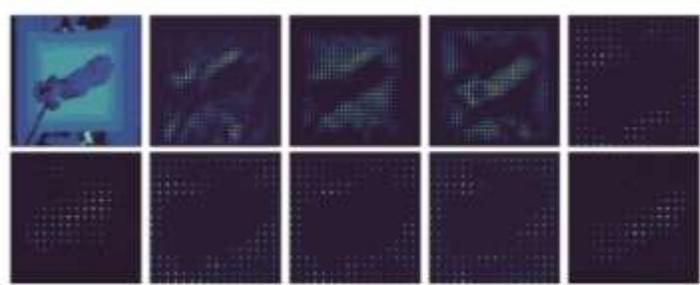


*Journal of Animal Science*, 2022, XX, 1–13  
<https://doi.org/10.1093/jas/skac242>  
 Advance access publication XX XX XXXX  
 Technology in Animal Science



## Predicting ribeye area and circularity in live calves through 3D image analyses of body surface

Joseph G. Caffarini,<sup>1,†,✉</sup> Tiago Bresolin,<sup>2</sup> and Joao R. R. Dorea<sup>1,†,1,✉</sup>



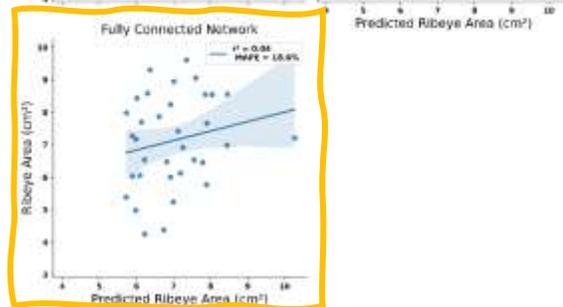
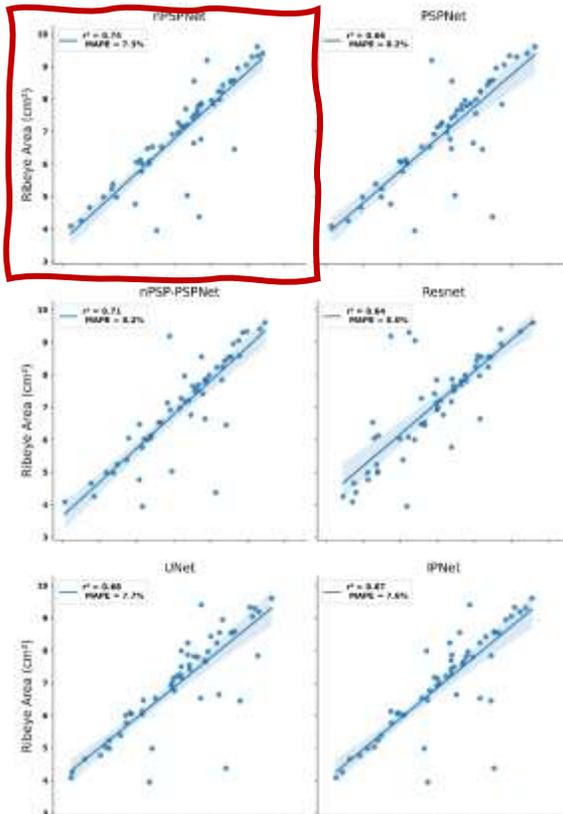
We evaluated 6 deep neural networks for segmentation:  
 nPSPNet, ResNet, Unet, IPNet, nPSP-PSPNet, PSPNet

# Predicting Ribeye Area and Shape of Live Calves

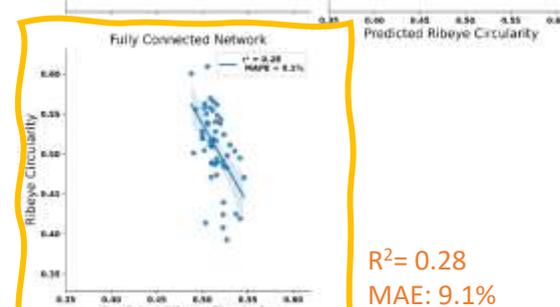
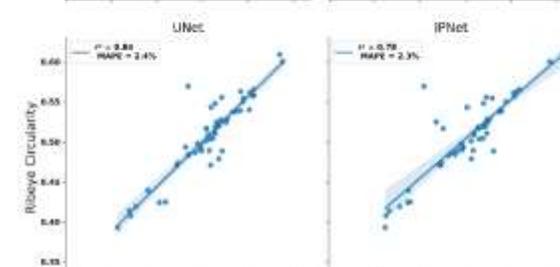
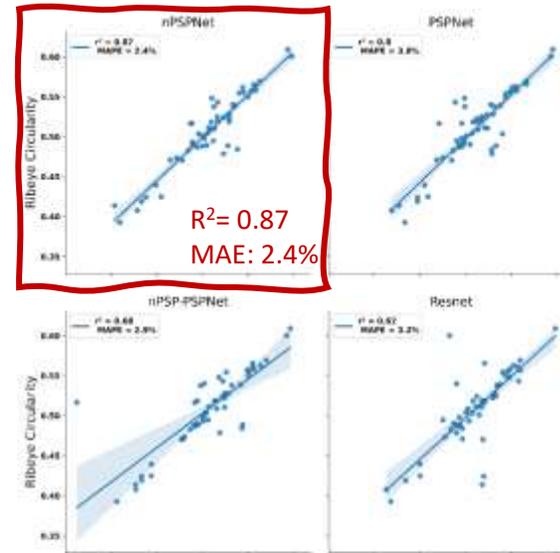


Testing set:

CNN Features  
 $R^2 = 0.74$   
 MAE: 7.3%

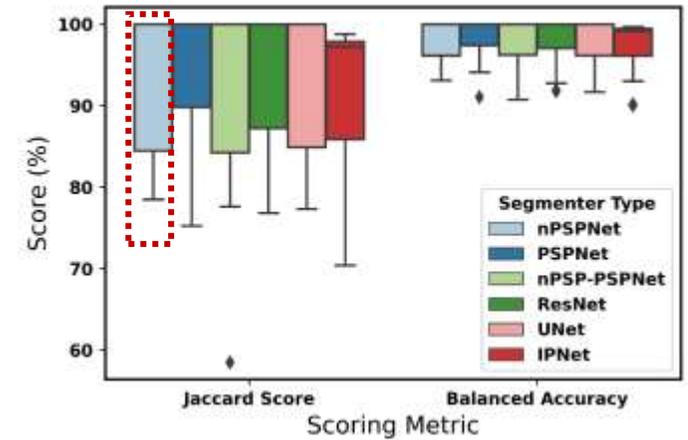


Biometric Features  
 $R^2 = 0.04$   
 MAE: 18.6%

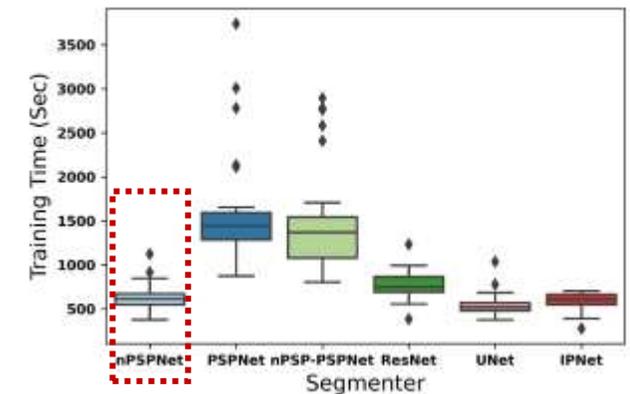


$R^2 = 0.28$   
 MAE: 9.1%

a



b

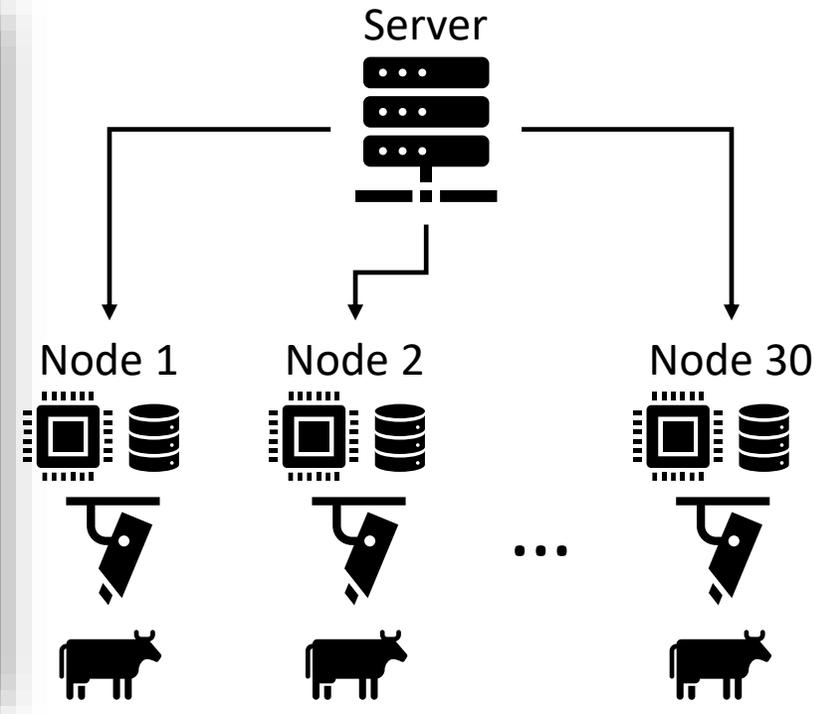
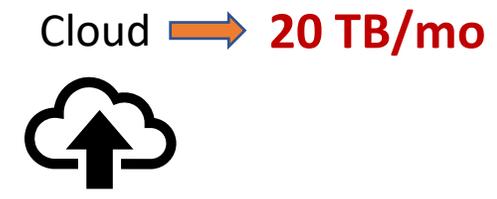


# Combining high-throughput phenotyping and genomic

## -Data Integration:

**Body growth** + Mammary gland development + Genomic information

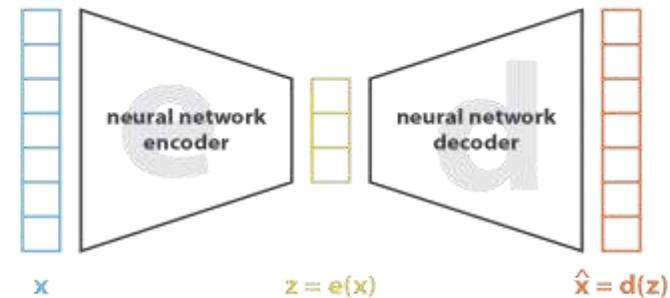
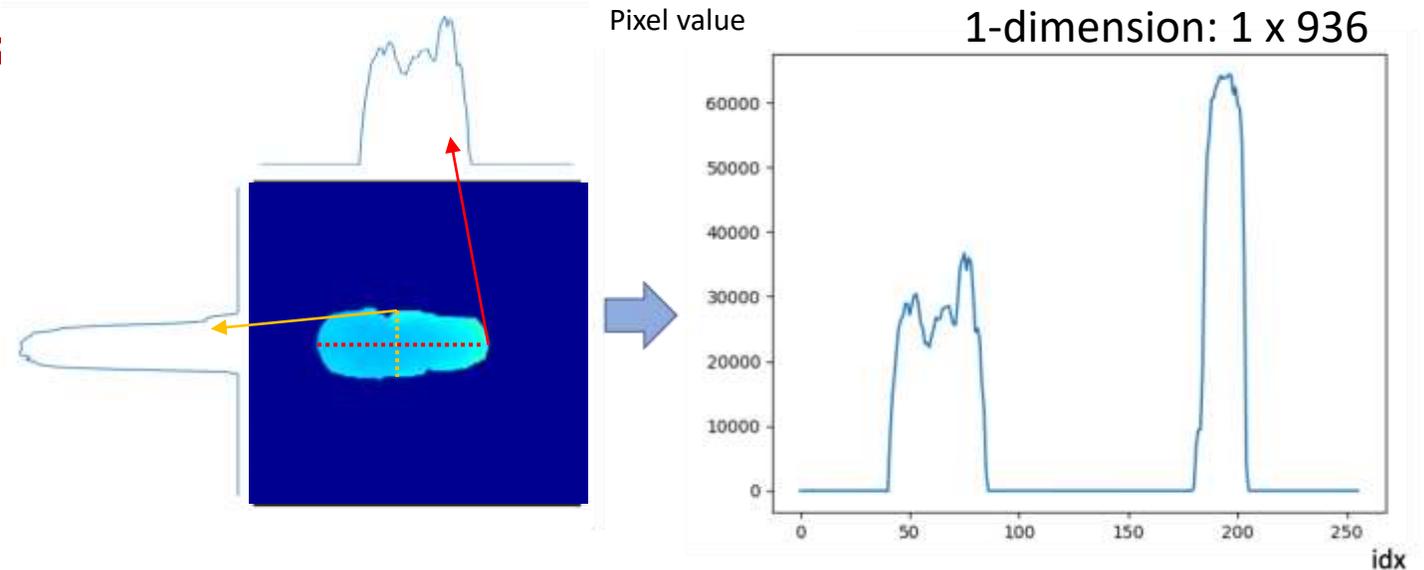
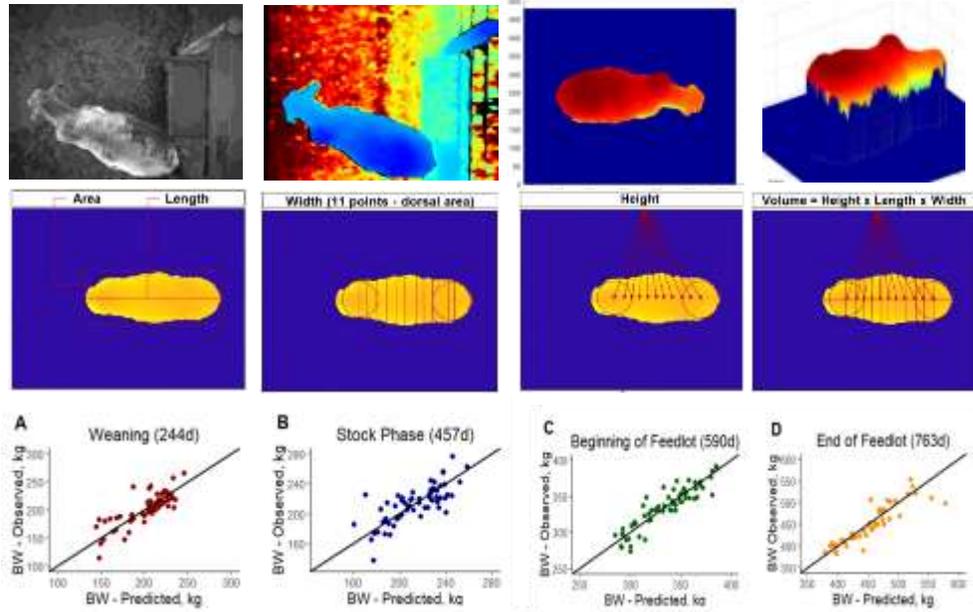
-From birth to first lactation (240 animals):



# We cannot transfer 20 TB of data/mo

## Reducing Data Dimension - Autoencoders

### Predicting Body Weight from 3D images



$$\text{loss} = \|x - \hat{x}\|^2 = \|x - d(z)\|^2 = \|x - d(e(x))\|^2$$

**Cattle:** Cominotte et al., 2020 –  
*Livestock Science* 232:103904

**Pigs:** Fernandes et al., 2019 –

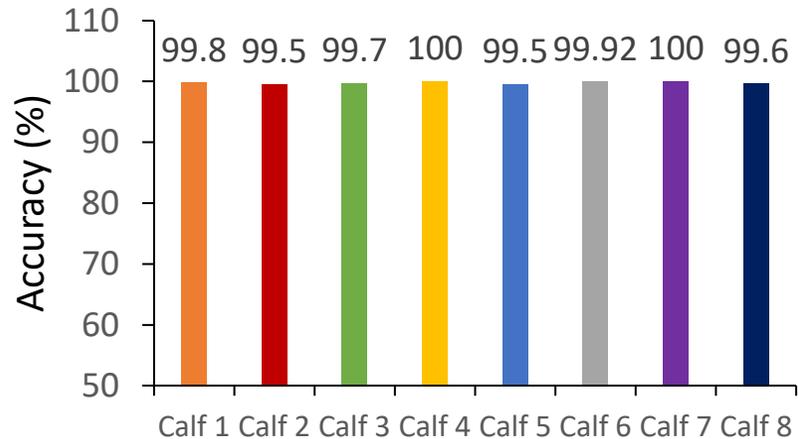
*Journal of Animal Science* 97:496-508

**Other groups**



# Predicting Feeding Behavior

- 1,546 images were used to train a deep learning algorithm for object detection (**YOLOv3**);
- 663 extra images were used for testing



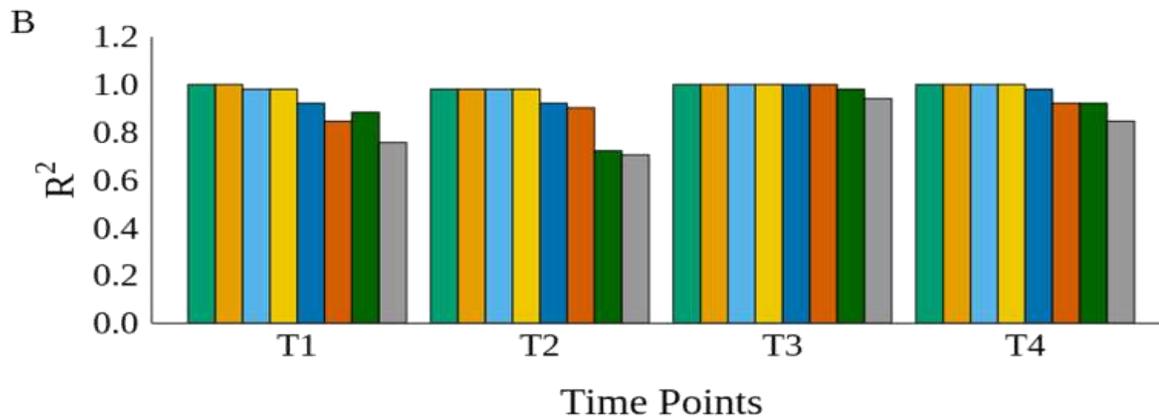
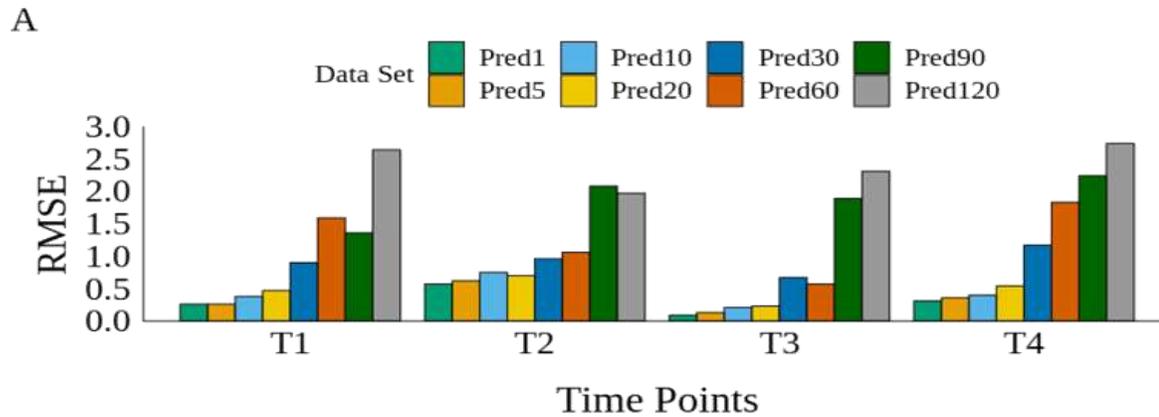
## The $R^2$ between observed and predicted:

- **Total eating time:** 0.99
- **Visit duration:** 0.77
- **Interval between visits:** 0.70
- **Visits:** 0.55



# Do we need to perform inference every second?

- Total time spend at the feed bunk – Inference performed in images sampled every 1, 5, 10, 20, 30, 60, 90, and 120 seconds.**



# Understand feeding behavior patterns in dairy cows

- 348 randomly selected images used to train Mask R-CNN
- Independent testing set: 711 images
- Cow mouth detection:
  - F1-score = 94.9%
  - Precision = 90.7%
  - Recall = 99.5%
  - NPV = 87.5%
  - PPV = 90.7%

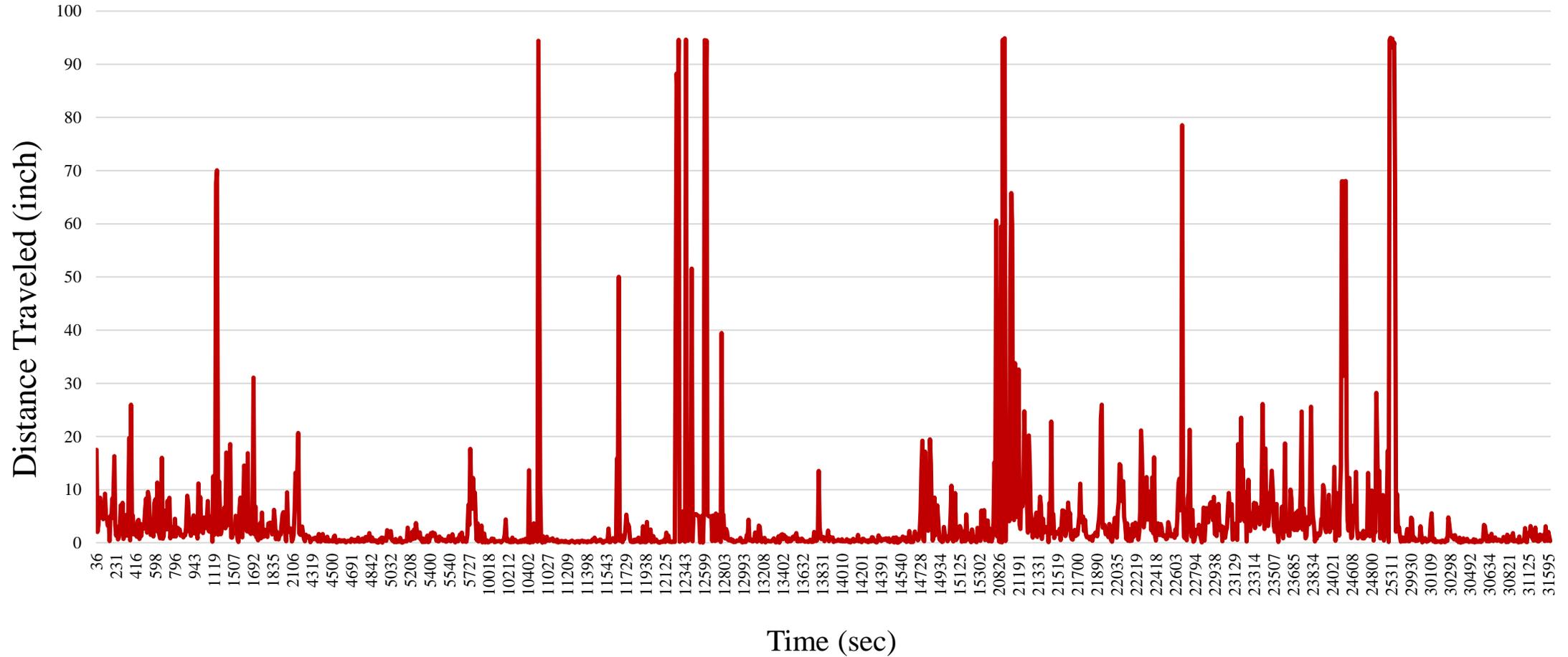


# Understand feeding behavior patterns in dairy cows

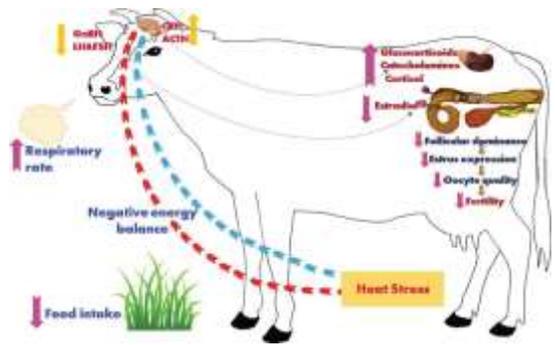
- **Calculated metrics:**
  - Distance traveled
  - Acceleration
  - Velocity
  - Feeding time
  - Standard deviation
- Anticipated metrics:
  - Number of meals
  - Meal frequency
  - Meal duration
  - Eating Time



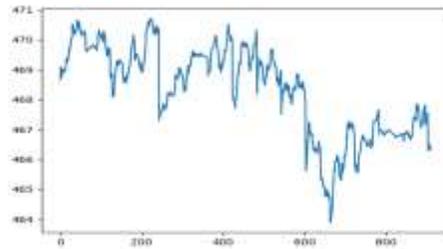
# Total Distance Traveled – x and y-axis



# Computer Vision System to Monitor Respiration Rate



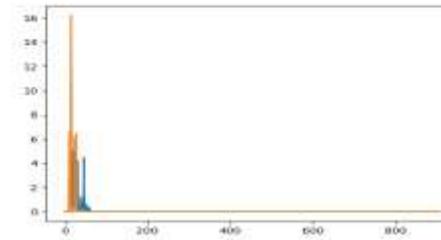
Pixel intensity  
(original domain)



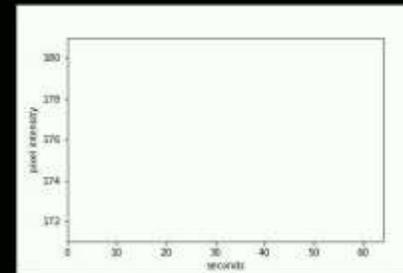
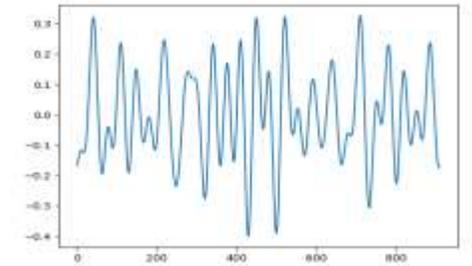
Fast  
Fourier  
Transform

$$x[k] = \sum_{n=0}^{N-1} x[n]e^{-j2\pi kn}$$

Power Spectra Density



Frequency domain  
(Fast Fourier Transform)

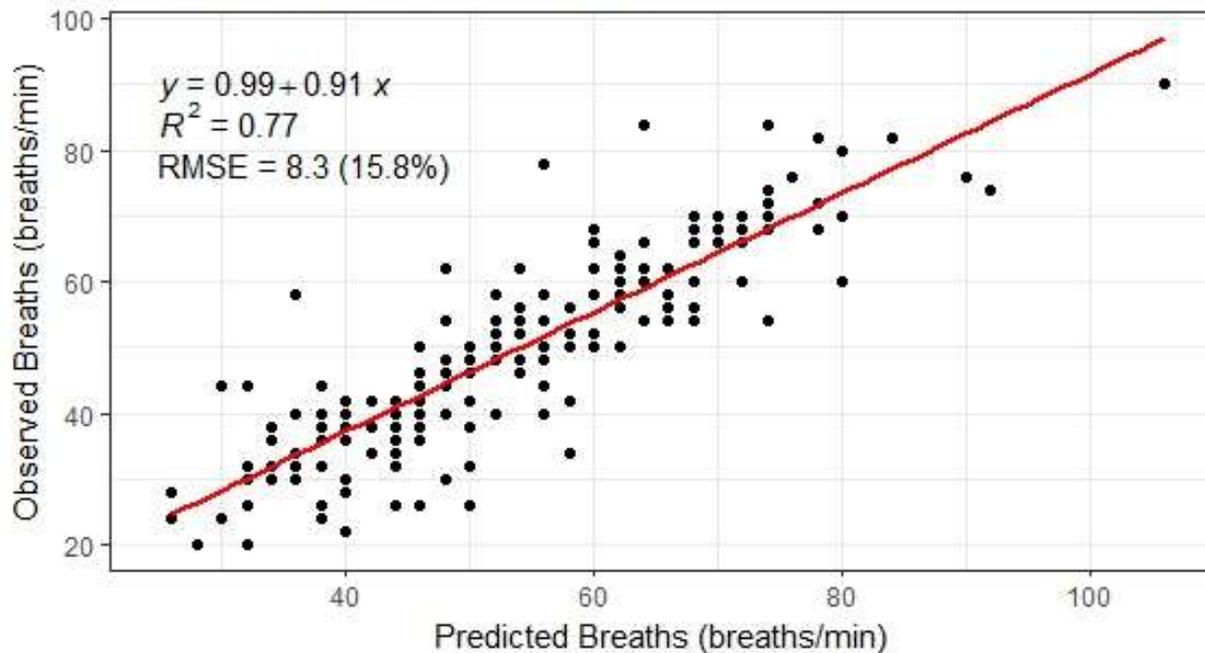


Blue cow:  
52 breaths/min

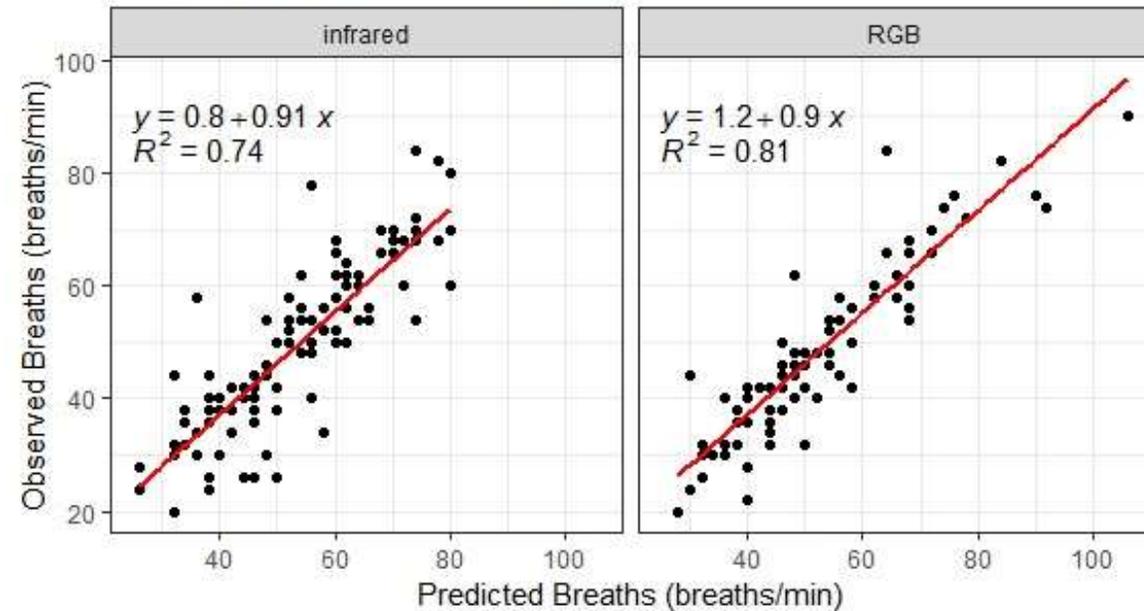
# Predictive Performance – Respiration Rate

- 168 videos (30-seconds segments) from 32 lactating cows
- Infrared images (night period)
- RGB images (day period)

Observed vs Predicted Breaths



Observed vs Predicted Breaths

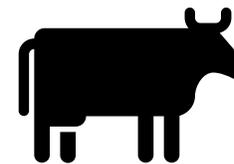


# Third-Person view



Animal

Human



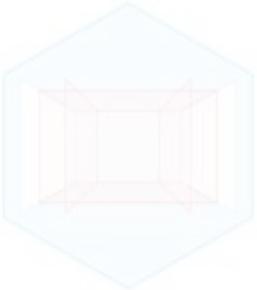
Processes

# V-A-M Reality Systems

## VR

Virtual Reality

Fully artificial environment



Full immersion in virtual environment



## AR

Augmented Reality

Virtual objects overlaid on real-world environment



The real world enhanced with digital objects



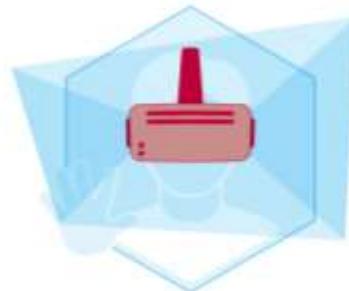
## MR

Mixed Reality

Virtual environment combined with real world



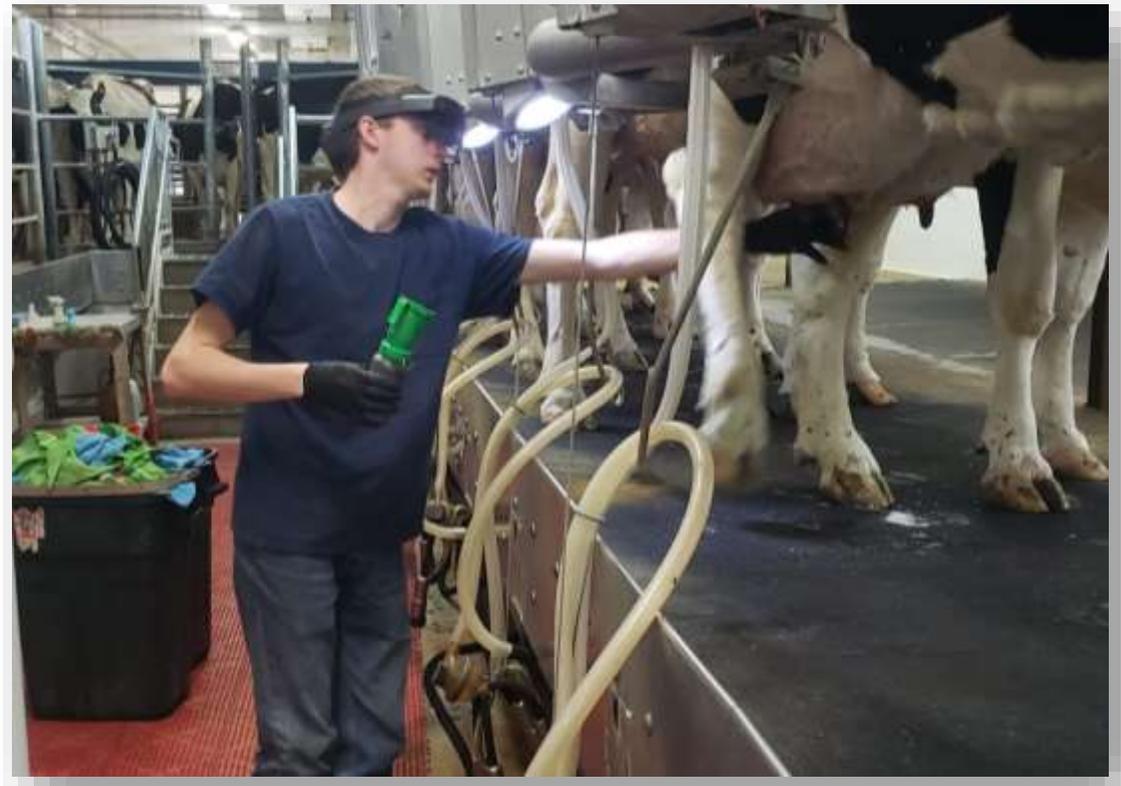
Interact with both the real world and the virtual environment



# From third- to first-person view



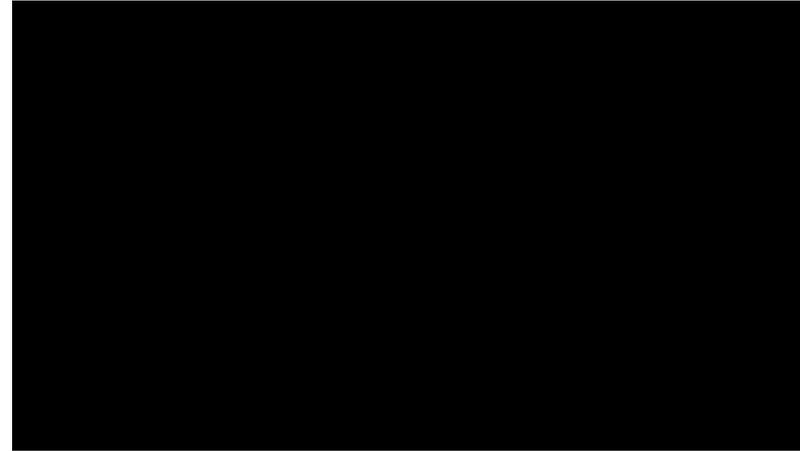
**Third-person view**



**First-person view**



# First-Person view: **processes** and **training**



**Unique tool** to train new employees

**Augment** human perception

**Track standard protocols (hygiene )**

**Human x Animal** interactions:

Pre-milking stimulation is associated with higher average milk flow rates and reduced unit attachment time

(1)Sagi, R., R. C. Gorewit, W. G. Merrill, and D. B. Wilson. 1980a. Premilking stimulation effects on milking performance and oxytocin and prolactin release in cows. J. Dairy Sci. 63:800–806. / (2)Rasmussen, M. D., E. S. Frimer, D. M. Galton, and L. G. Petersson. 1992. The influence of premilking teat preparation and attachment delay on milk yield and milking performance. J. Dairy Sci. 75:2131–2141. / (3) Gorewit RC, Gassman KB. Effects of duration of udder stimulation on milking dynamics and oxytocin release. J Dairy Sci. 1985 Jul;68(7):1813-8

# Acknowledgments



United States Department of Agriculture  
National Institute of Food and Agriculture



# Final Considerations

---

- **Digital technologies** are **crucial** to collect cheaper, precise, and real-time phenotypes
- **Animal-level** information is a **very important component** of any integrated databases
- Leverage **Artificial Intelligence** Systems: Applications in Livestock (**Dairy and Beef**)
- It is not about new questions only! It is about **unanswered questions!**
- **Digital Agriculture**: undergrad and grad courses (livestock, crop, water, soil - data management, storage, and analyses – cloud computing)
- **New generation** of students/professionals
- **Multidisciplinary teams**: Collaboration across campus



Thank you!

 @digitallivestock

 @jrrdorea

joao.dorea@wisc.edu