Innominds

ENGINEERING INNOVATION

AI for Agriculture and the Connected Farm

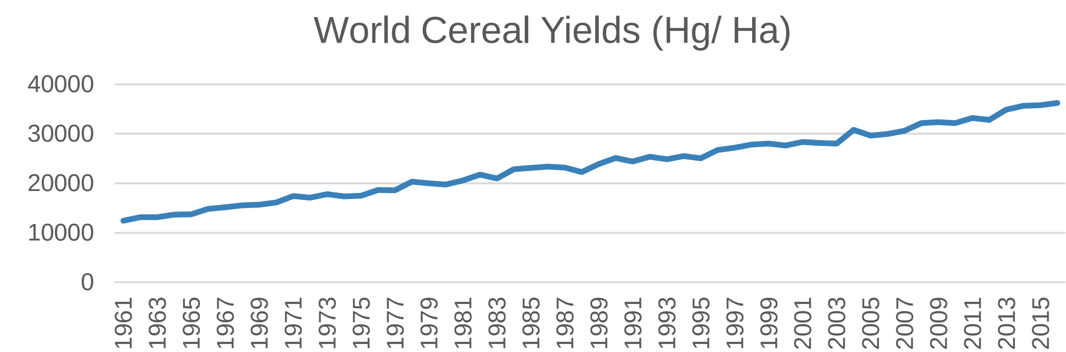
January 19, 2018

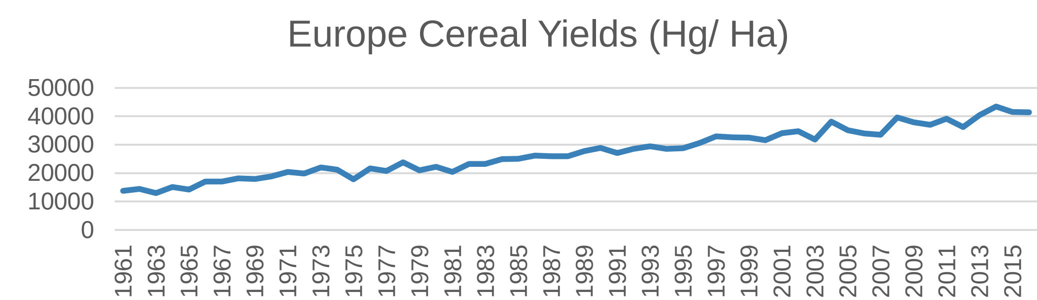
AGENDA

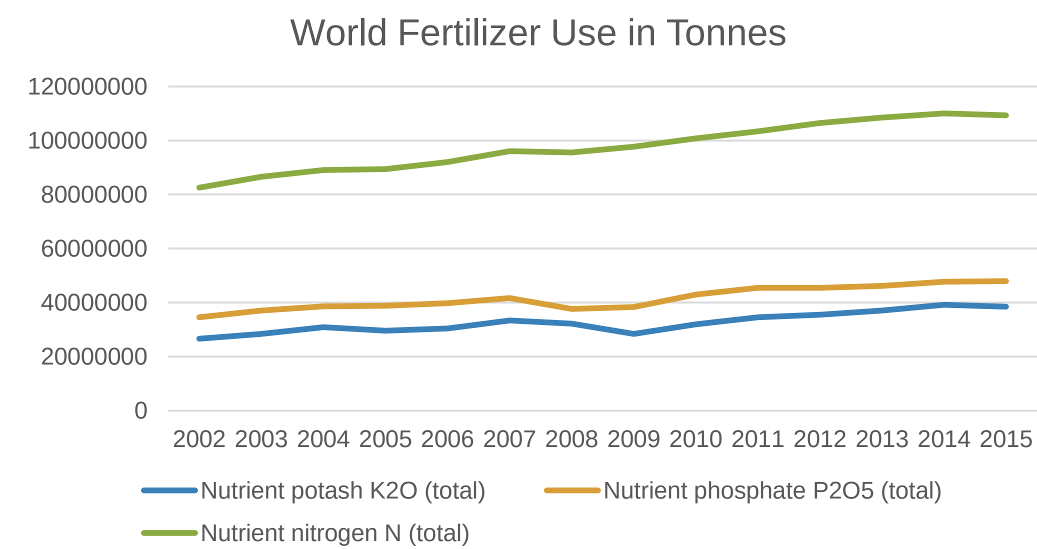
1. Agriculture – Challenges, Opportunities and the Future
2. IoT and the Connected Farm
3. AI for Geo Spatial Analysis of the Soil Attributes
4. Case study – Canadian Farms
5. Deep Learning for Geo-Spatial Analytics

Agriculture – Challenges and Opportunities

Source: FAO







- Peak Farmland to be reached very soon and no major increases in arable land visible
- Global Yields rising but growth mostly due to new irrigated areas, new seeds and varieties in emerging economies and through increased use of Inputs
- New Agricultural revolution necessary to make it more efficient and sustainable
- In India, farming is in distress characterized by higher costs (labor and inputs), no major increase in yields and insufficient prices despite a high food inflation
- Digital Farming leading to Precision Farming is an opportunity and necessity to bring about a Food Revolution

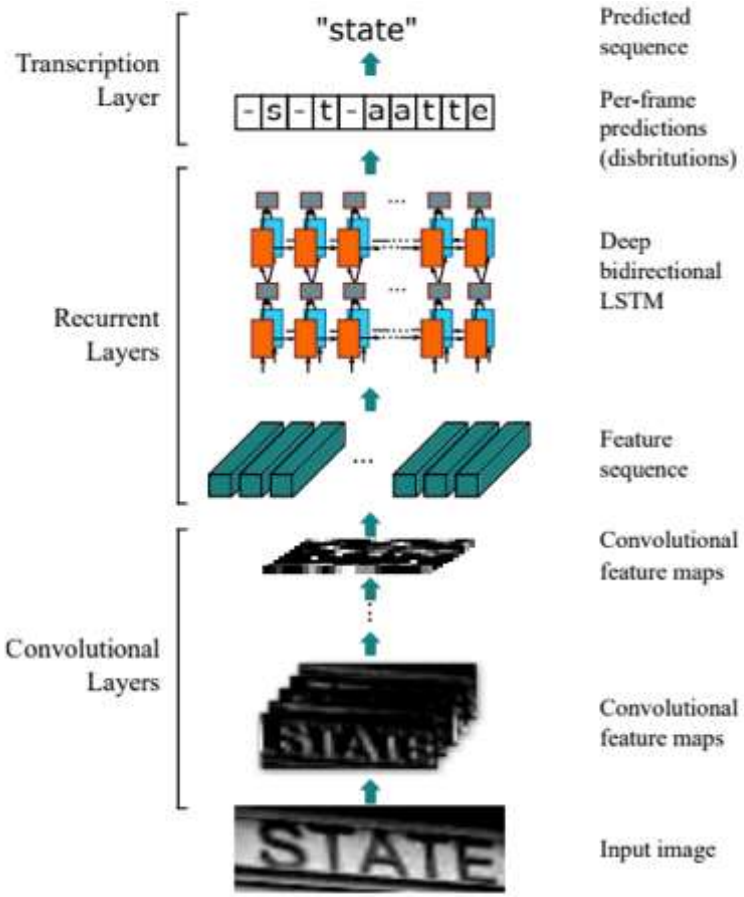
Digital Farming – The Connected Farm and IoT

- Connecting the farm for
  - Soil attributes
    - Physical
    - Chemical
    - Topological
  - Weather
  - Crop health
  - Water and other resources
  - Machinery
- Enabling automation with data from sensors and images backed by computer vision
- Benefits
  - Matching crop varieties and sowing strategies to soil
  - Controlling water usage for optimal plant growth.
  - Determining custom fertilizer profiles based on soil chemistry.
  - Determining the optimal time to plant and harvest.
  - Reporting weather conditions.

Image Source: Bayer Research

AI for the Connected Farm

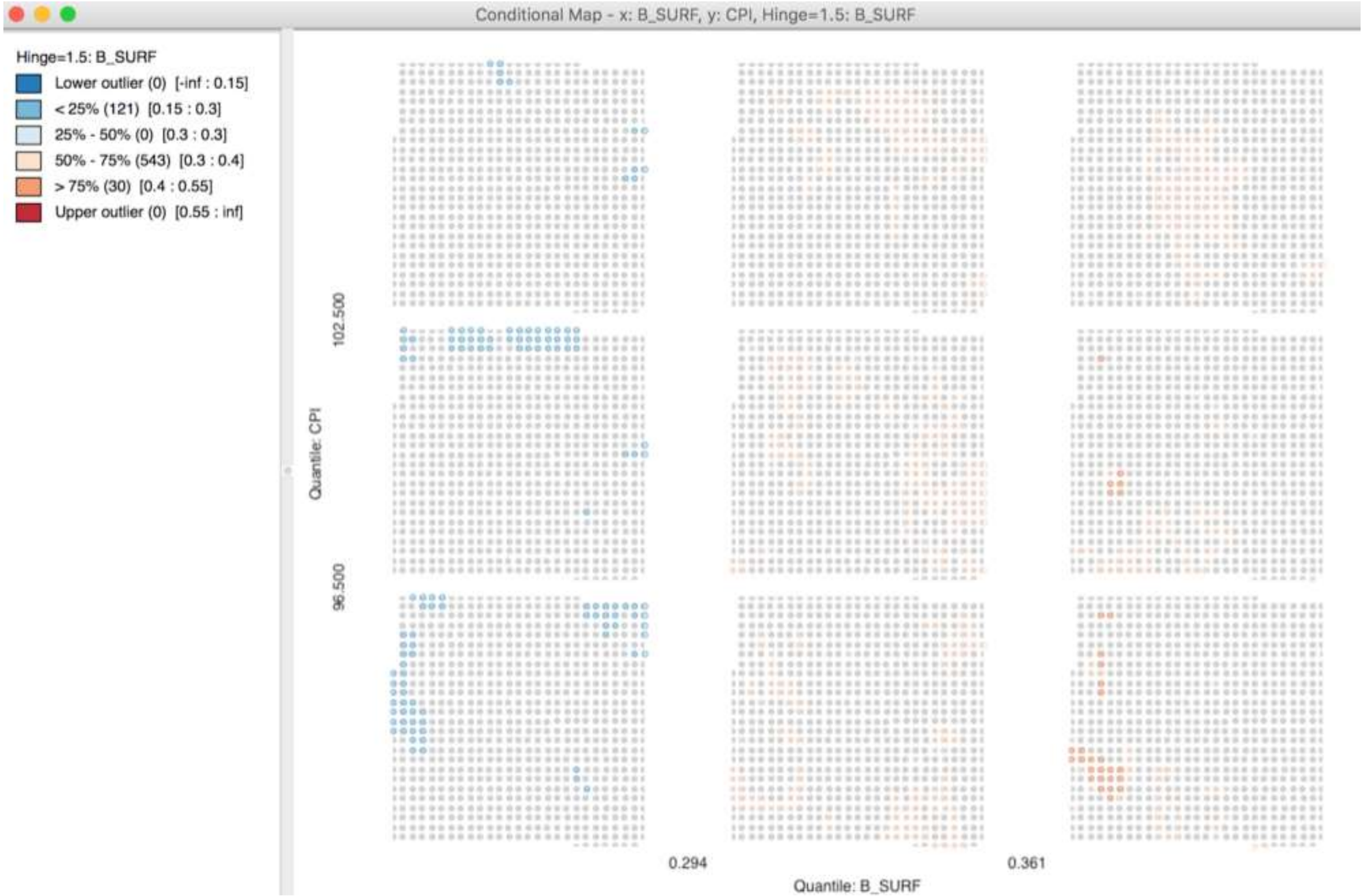
- Taking Research from inside the lab or on controlled zones to real time and large scale farms with the combination of IoT + Big Data + Cloud + Machine Learning
- Early Detection and Protection with pattern recognition and interaction effects not easily detectable otherwise – hence the need for Deep Learning
- Objective to increase Yields, reduce costs and protect the environment
- Use Cases
  - When to sow
  - What to sow
  - Harvest timing
  - Controlled use of fertilizer, pesticide and inputs
- Combination of structured – unstructured and Temporal & Geo-Spatial dimension data harder to model STARMA
- Big Data such as Spark with its Spatial Extensions to SQL and Time Series functions enabling this change



Case Study – Farm Lands in Saskatchewan, Canada

- 5 Farm land of data of Chickpea
- 78 Soil attributes with crop productivity index (CPI) and yields
- Historic attribute values and yields for the last 3 years
- Traditional techniques
  - Mutual Information
  - Recursive Feature Extraction
  - Random Forest
  - PCA
  - Select K Best Method
  - Decision Tree Regressor/ Extra Tree Regressor
  - Spatial Lag
  - Spatial Correlation
- Deep Learning
  - Segment based learning
  - Sequential learning

Spatial Correlation



Spatial Correlation

REGRESSION DIAGNOSTICS

MULTICOLLINEARITY CONDITION NUMBER	546722.585

TEST ON NORMALITY OF ERRORS

TEST	DF	VALUE	PROB
Jarque-Bera	2	16.685	0.0002

DIAGNOSTICS FOR HETEROSKEDASTICITY

RANDOM COEFFICIENTS

TEST	DF	VALUE	PROB
Breusch-Pagan test	77	214.715	0.0000
Koenker-Bassett test	77	162.381	0.0000

DIAGNOSTICS FOR SPATIAL DEPENDENCE

TEST	MI/DF	VALUE	PROB
Moran's I (error)	0.5612	27.030	0.0000
Lagrange Multiplier (lag)	1	751.670	0.0000
Robust LM (lag)	1	164.229	0.0000
Lagrange Multiplier (error)	1	594.824	0.0000
Robust LM (error)	1	7.383	0.0066
Lagrange Multiplier (SARMA)	2	759.053	0.0000

Interpretation of the model results

- Spatial co-relation present in the data
- Multi-colinearity found very high among the features
- Test on normality of errors was violated using Jarque-Bera test
- Spatial dependency were found from Moran's I error test
- Based on the above tests results traditional technique provides poor results (MI, RF)
- Robust LM test indicates Spatial Lag model for better results
- Spatial Lag model seems to be better based on below compare to Error model
    - Pseudo R-squared
    - Log Likelihood
    - Akaike Info Criterion
    - Schewarz creterion
- Top features were considered which has high magnitude and P values less than 0.05

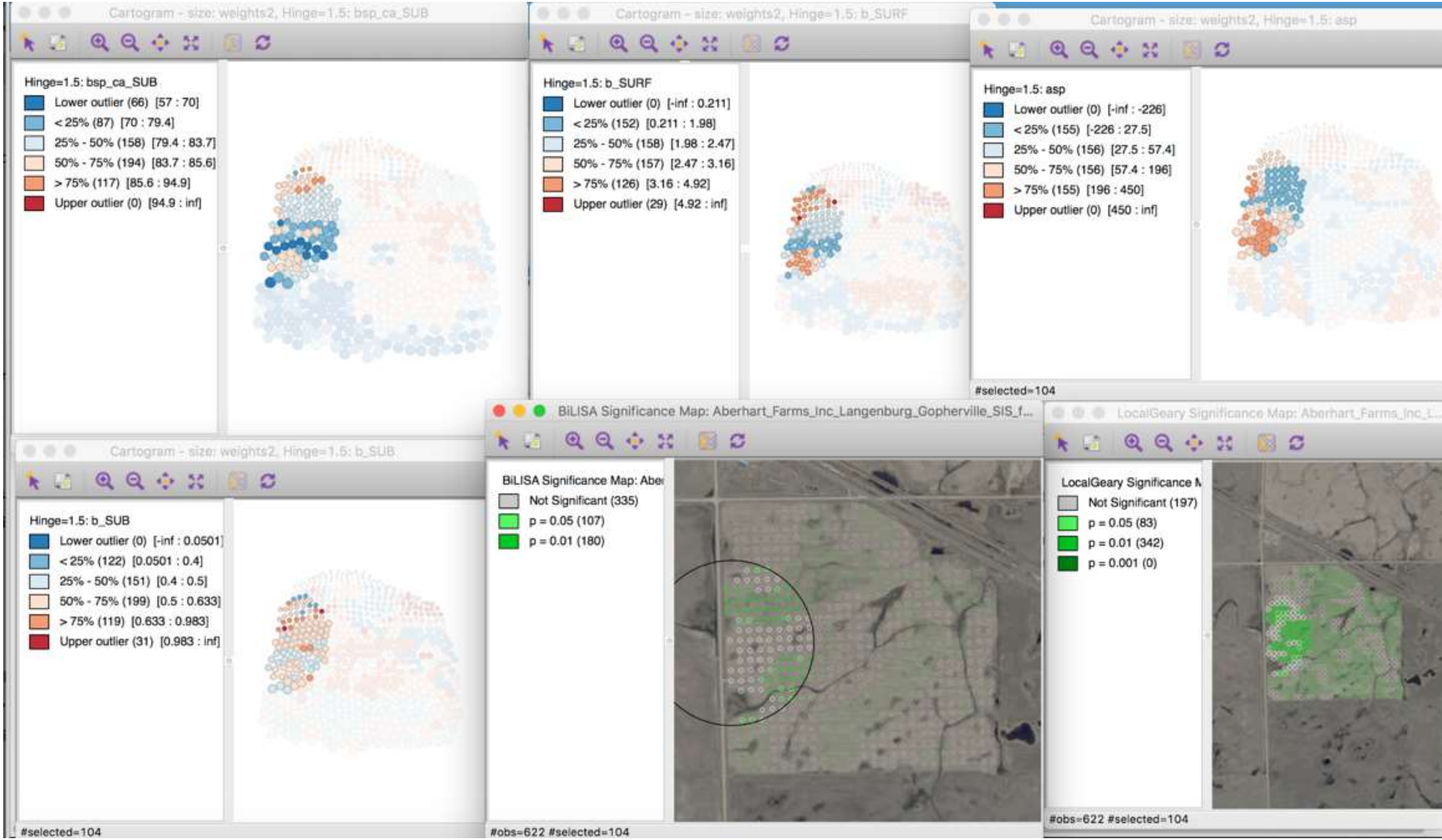
Summary Results

New Data - Saskatchewan			2 Category (Low and High)				
Farmland	Mutual Information	Random Forest		Spatial Lag		Spatial Error	Error
Aberhart_Farms_Inc_Langenburg_Gopherville_SIS_full.shp	dem emv rlsat_SUMM dprr_SUMM bsp_na_SUR RMSE == 5.34 R2 == 0.40	dem emv rzpwp_SUMM paw30_SUMM b_SURF	Confusion Matrix [[11 21] [5 110]] AUC Score 0.650135869565 precision recall f1-score support Medium 0.69 0.34 0.46 32 High 0.84 0.96 0.89 115 avg / total 0.81 0.82 0.80 147 RMSE == 5.31359740774 R2 == 0.391394433001	SUMMARY OF OUTPUT: MAXIMUM LIKELIHOOD SPATIAL LAG (METHOD = ORD) -- Data set : SPATIAL LAG Weights matrix : Kernel Dependent Variable : CPI Number of Observations: 581 Mean dependent var : 4.6012 Number of Variables : 81 S.D. dependent var : 0.0695 Degrees of Freedom : 500 Pseudo R-squared : 0.8901 Spatial Pseudo R-squared: 0.5574 Sigma-square ML : 0.001 Log likelihood : 1289.076 S.E of regression : 0.023 Akaike info criterion : -2416.153 Schwarz criterion : -2062.608	b_SURF b_SUB asp bsp_ca_SUB	SUMMARY OF OUTPUT: MAXIMUM LIKELIHOOD SPATIAL ERROR (METHOD = FULL) -- Data set :Spatial Error Weights matrix : Kernel Dependent Variable : CPI Number of Observations: 581 Mean dependent var : 4.6012 Number of Variables : 80 S.D. dependent var : 0.0695 Degrees of Freedom : 501 Pseudo R-squared : 0.3454 Sigma-square ML : 0.000 Log likelihood : 1298.860 S.E of regression : 0.022 Akaike info criterion : -2437.720 Schwarz criterion : -2088.540	bsp_k_SUB bsp_ca_SUR bsp_ca_SUB b_SUB asp

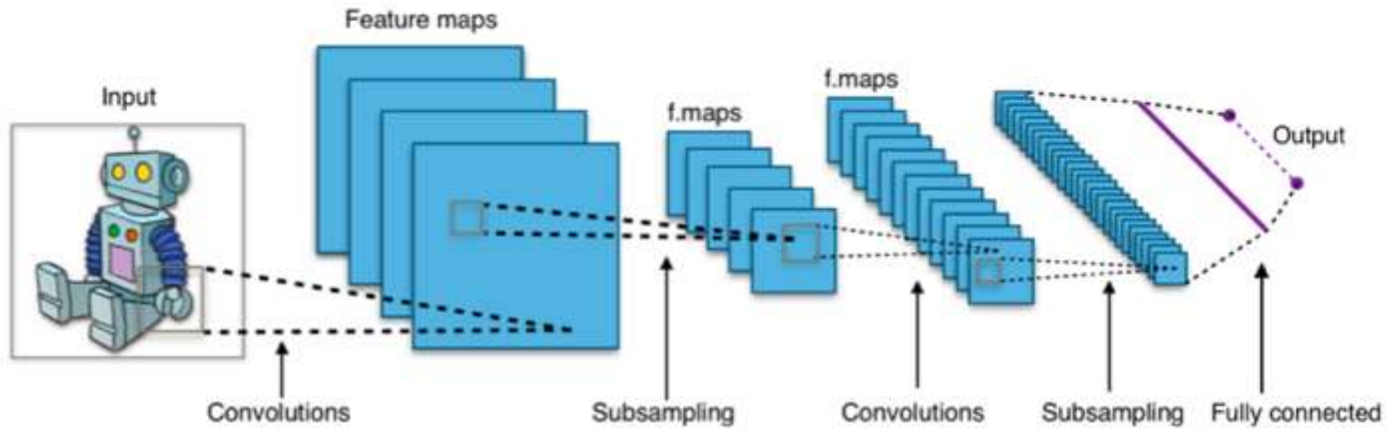
Spatial Regression Results

- b_SURF
- b_SUB
- asp
- bsp_ca_SUB

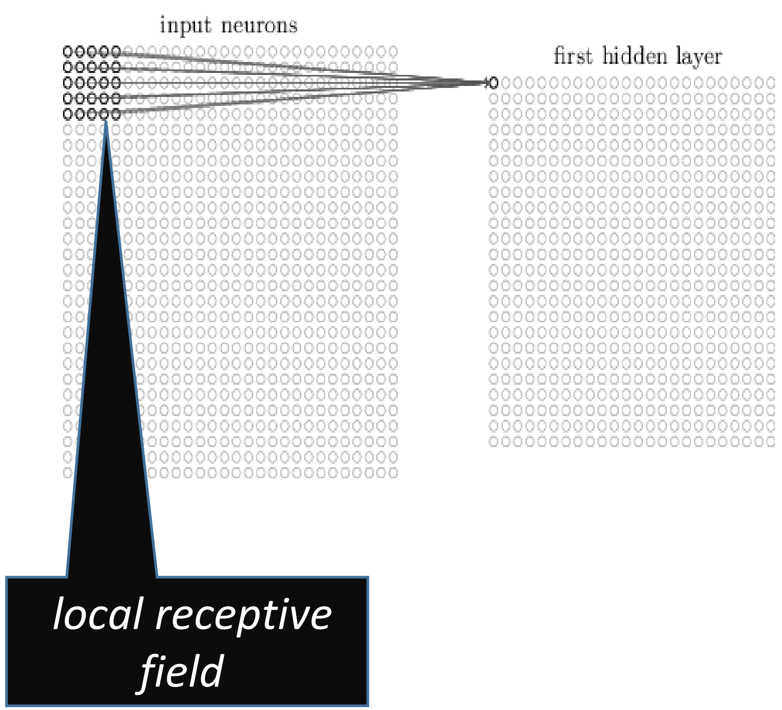
Actionable Insights

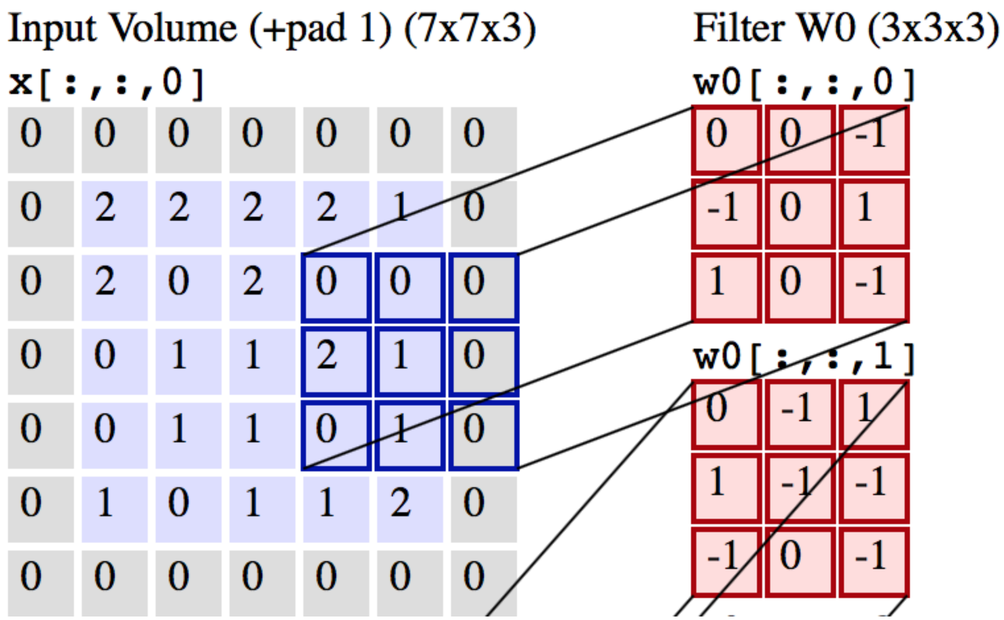


Deep Learning – CNN for Segment Learning

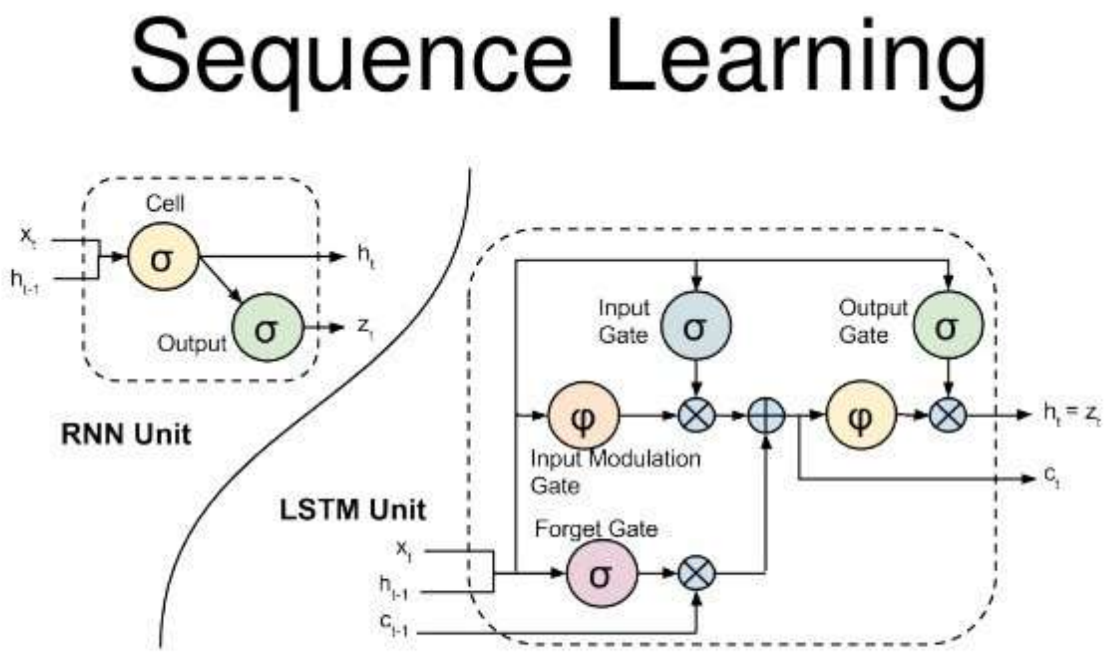


- Input Layer
- Convolutional Layer
- Pooling Layer
- Fully Connected Layer



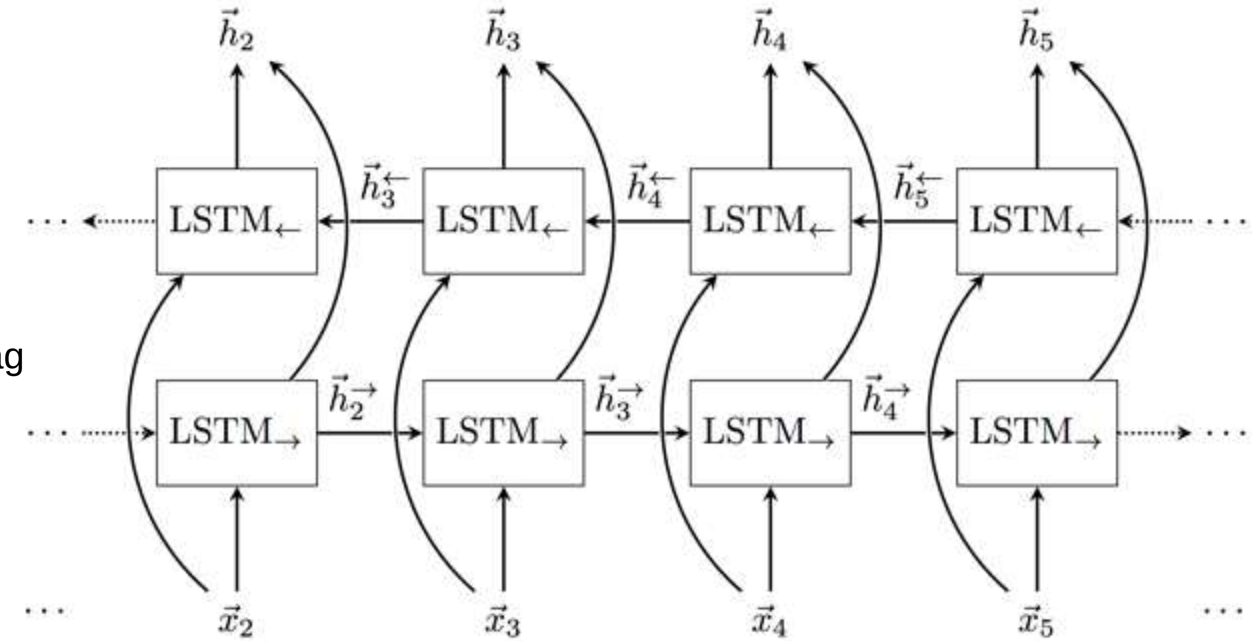


Deep Learning – RNN for Sequence Learning

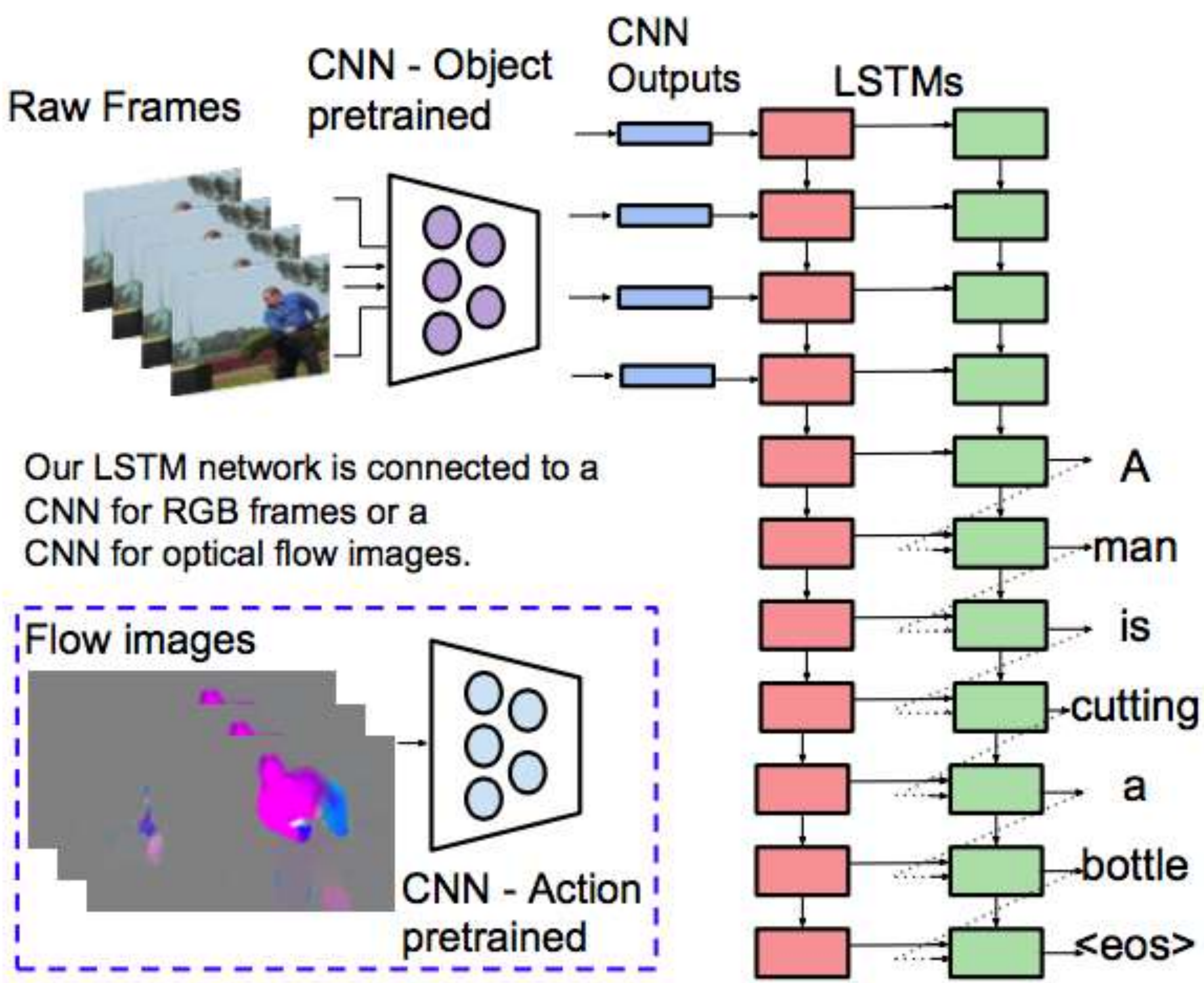


- Memory Cell
- Input Gate
- Output Gate
- Forget Gate

Bidirectional LSTMs allow to go back as well to detect lead and lag effects



Combining Space Time Correlations

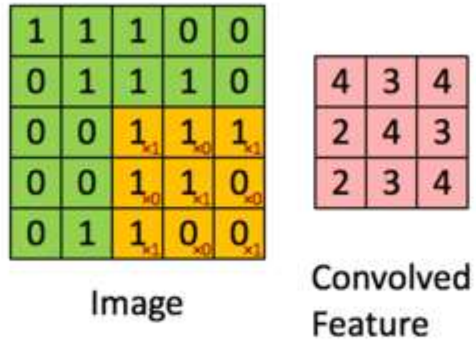


Spatial Data can be combined effectively with Time Series Data through stacked network

Keeping Yield of a portion of the farm as the class or continuous output

CNNs first address the spatial correlation through

1. Channel (all the 78 attributes instead of the regular R, G and B used in image classification)
2. Convolutions reduce the Space dimension through trained Sparse Autoencoders



3. Provide reduced temporal class outputs from CNN to an LSTM for sequence prediction
4. Feature importance and contribution extracted through techniques by skip feature, introducing noise and by using partial connected layers

THANK YOU

We love to innovate

We like creating products

Happy when we all succeed

GLOBAL **LOCATIONS**

NORTH AMERICA

SAN JOSE (HQ)

2055 Junction Avenue,
Suite 122,
San Jose, CA 95131.
Phone: +1 408-434-MIND
Fax: +1 408-434-7061

EUROPE

SWITZERLAND

Innominds GmBH,
Gotthardstrasse 28,
CH-6304 Zug,
Switzerland

INDIA

HYDERABAD

4th Floor, Waverock SEZ, Road #2, Financial District
Hyderabad – 500 032
Telangana
Phone: +91 (0)40 66126300

HYDERABAD

Kala Jyothi building,
Survey #185, Kondapur
Hyderabad – 500 133
Telangana
Phone: +91 (0)40 46126700

BANGALORE

3rd floor, Anand Bhavan
#9, Netkallappa circle, Basavanagudi
Bangalore – 560 004
Karnataka
Phone: +91 (0)80 40447400

VISAKHAPATNAM

Hill No 3 APIIC IT / ITES SEZ,
Rushikonda
Visakhapatnam -530034
Andhra Pradesh