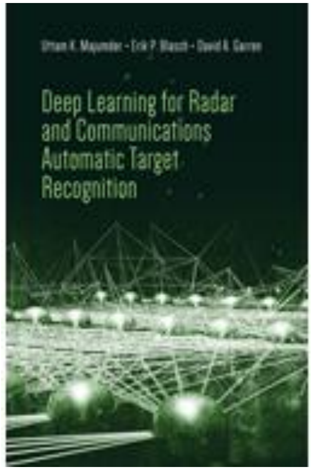
Machine Learning Techniques for Radar Automatic Target Recognition (ATR)

Uttam K. Majumder

Chapter 9:
RF ATR Performance Evaluation

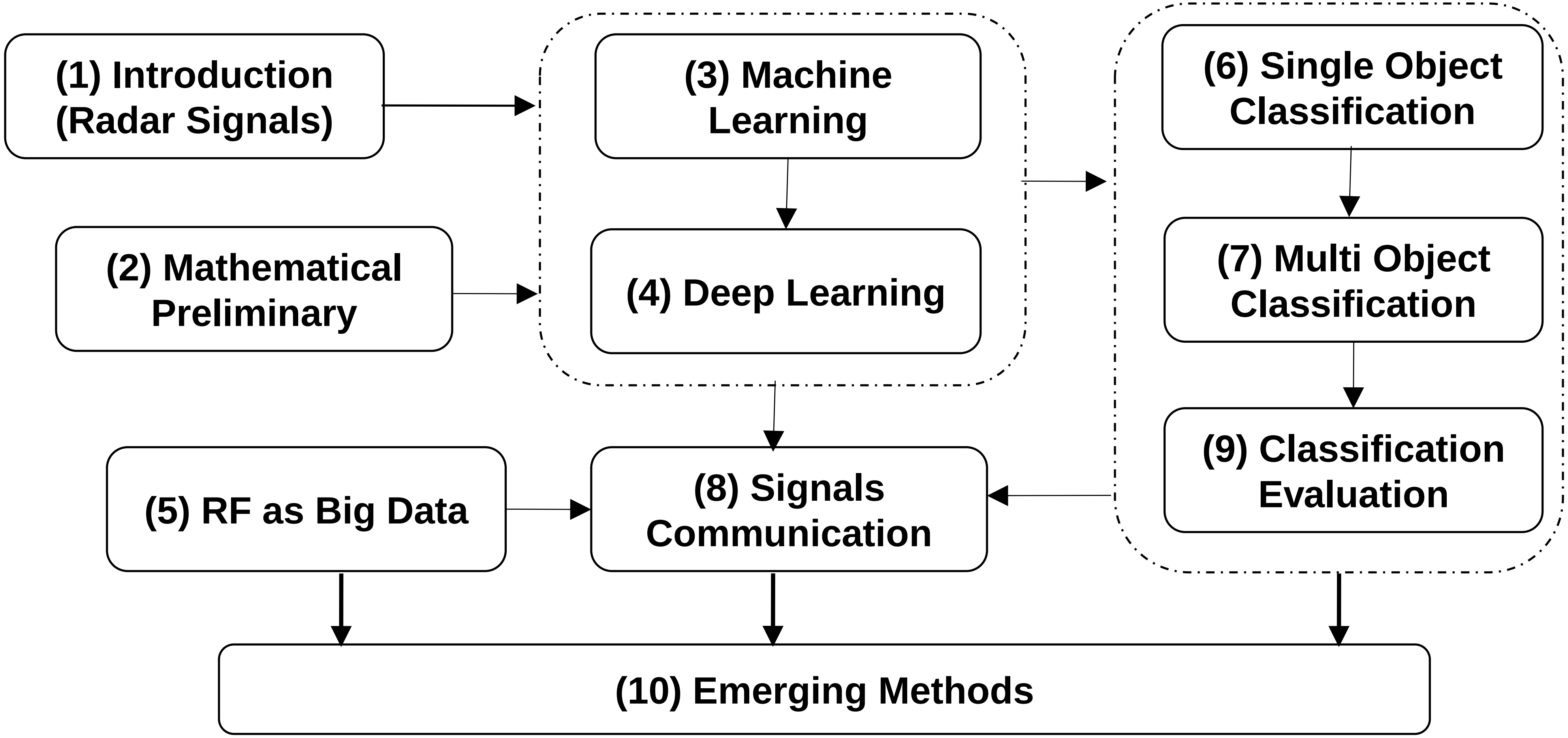
Lecture Outline

1. **Radio Frequency ATR: Past, Present, and Future: 30 min**
2. **Mathematics for Machine Learning / Deep Learning: 20 min**
3. **Review of ML Algorithms: 30 min**
4. **Deep Learning Algorithms: 30 min**
5. **RF Data for ML Research: 15 min**
6. **DL for Single Target Classification: 20 min**
7. **DL for Many Targets Classification: 20 min**
8. **RF Signals Classification: 20 min**
9. **RF ATR Performance Evaluation: 25 min**
10. **Emerging ML Algorithms for RF ATR: 30 min**



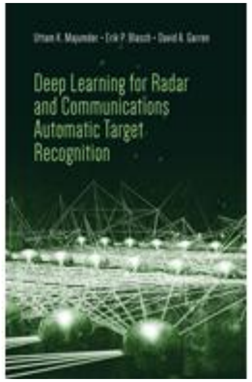
© Majumder, Blasch, Garren

Machine Learning Techniques for Radar Automatic Target Recognition (ATR)



© Majumder, Blasch, Garren

Chapter 9 Outline



© Majumder, Blasch, Garren

Data Information Fusion Group Model

E. Blasch, "Situation, Impact, and User Refinement," *Proc. of SPIE,* Vol. 5096,2003

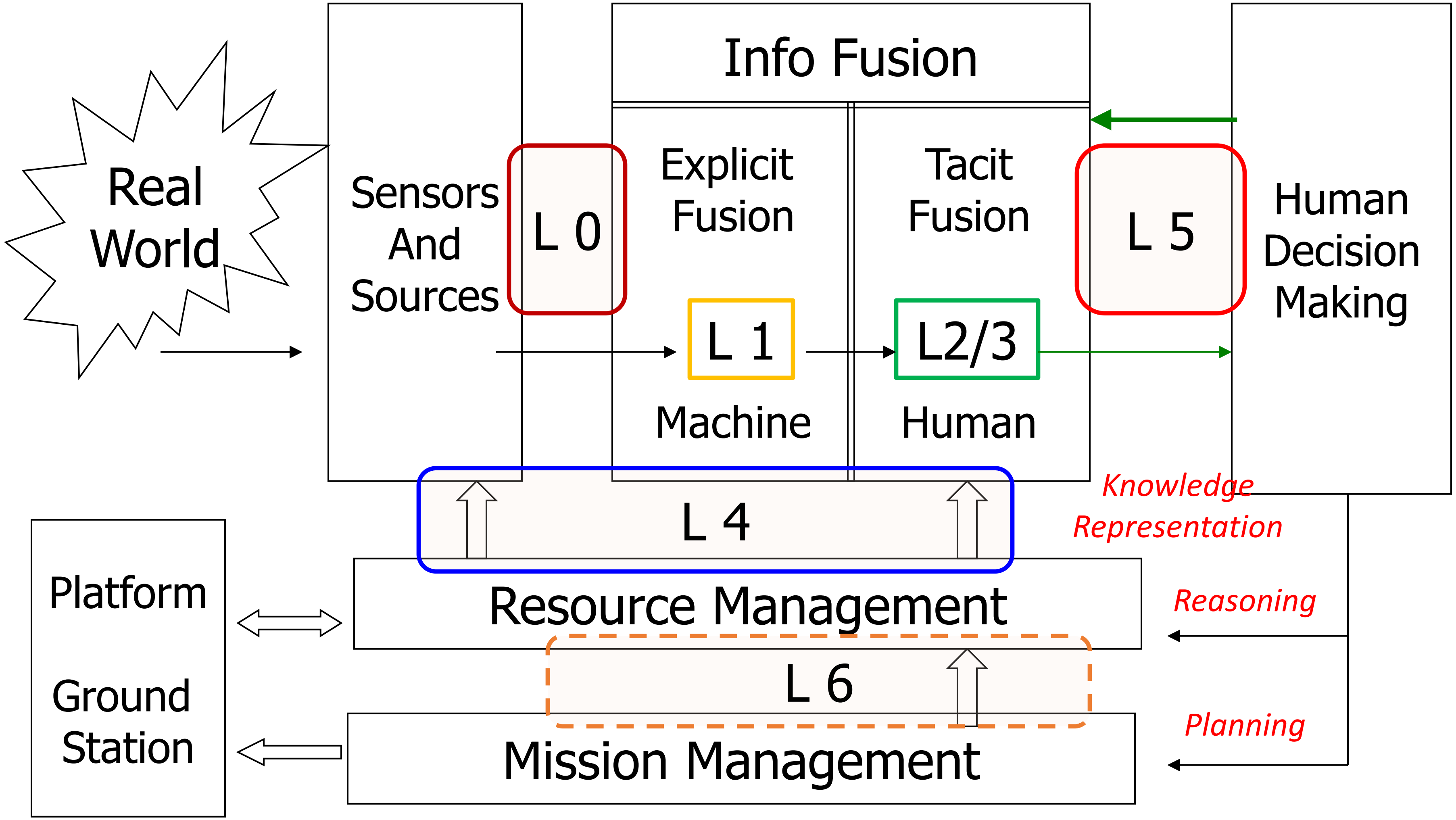
Figure 9.1

Levels (L)

Low-Level
ASSESSMENT
L1 – Object (ATR)

High-Level
ASSESSMENT
L2 – Situation
L3 – Threat

REFINEMENT
L4 – Process
L5 – User
L6 – Mission



E. Blasch, A. J. Aved, "Physics-Based and Human-Derived Information Fusion Video Activity Analysis," *Int'l. Conf. on Information Fusion*, 2018.

© Majumder, Blasch, Garren

Data Fusion Methods(1)

Automatic Target Recognition : Uses Data Fusion

Fusion Type	Description	Benefit	Limitation
Boolean Logic	Mathematical (Crisp) Rules of AND /OR / NOT/NOR	• Well Defined • Simple to implement • Expandable for large systems	• High uncertainty • Requires many rules • Reduces meaning and interpretation
Fuzzy Logic	Provides overlapping (soft) many-valued sets to process partial truth decision rules	• Handles imprecise information • Defines bounds for classification • Amendable to human analysis	• Not accurate • Decision rules are arbitrary • Does not scale to complex situations
Markov Chains	Probabilistic modeling of dynamic systemscomprised of defined finite states and transitions between states.	• Incorporate uncertainty of both the states and the model. • Computational complexity is minimal. • Offers scalabilty in complex scenarios	• Combinatoric explosion of states can reduce utility • Requires aggregation of states and pruning of superfluous stales • Calls for knowledge engineering to determine the meaning
Bayesian Networks	Probabilistic reasoning network for partial information providing a system of beliefs in given states that impact all other states in the network	• Models a priori conditional situations • Propagates information through the network methodically • Incorporates partial knowledge for uncertainty analysis	• Structure is static and cannot be changed adaptively. • Approximates the Interconnection functions (e.g. Gaussian sums) • Numerical approximations can have low bounds on the induced error

© Majumder, Blasch, Garren

Data Fusion Methods (2)

Automatic Target Recognition : Uses Data Fusion

Fusion Type	Description	Benefit	Limitation
Bayesian Networks	Probabilistic reasoning network for partial information providing a system of beliefs in given states that impact all other states in the network	• Models a priori conditional situations • Propagates information through the network methodically • Incorporates partial knowledge for uncertainty analysis	• Structure is static and cannot be changed adaptively. • Approximates the Interconnection functions (e.g. Gaussian sums) • Numerical approximations can have low bounds on the induced error
Entropy	Measures data information contained in a signal.	• Monitors and predictsmeasurement uncertainty • Handles discrete and continuous disparate sourcesuseful for data fusion • Widely adopted	• Entropy is unitless and can easily misinterpreted • Provides a relative assessment versus and absolute analysis • Uncertainty level is not conducive to high credibility
Dempster Shafer	Bayesian evidence accrual system that determines possibility (belief and plausibility) based on observations	• Incorporatespositive, negative, and conflicting information • Does not need a fully specified model for processing • Captures multi-decision reasoning	• Complexity of mathematical implementation does not scale • Requires approximation techniques with many fusion rules • Challenged with partial information

* Recent(2019): Deep Learning Evidential NN (replaces softmax with evidential reasoning)

© Majumder, Blasch, Garren

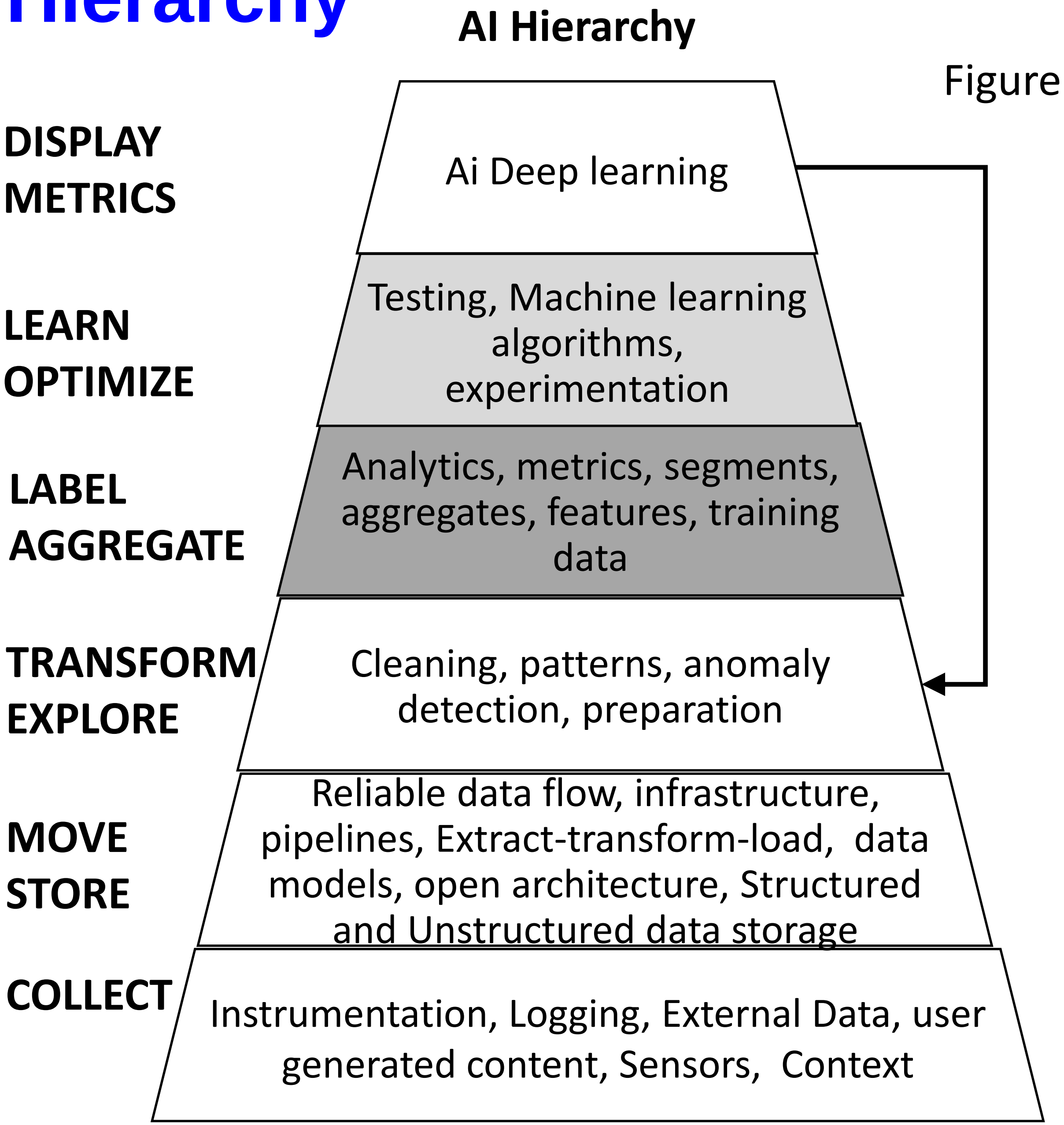
AI/ATR Hierarchy

Automatic Target Recognition Standards (ICD 203)

1. Sourcing
2. Uncertainty Analysis
3. Distinguish Judgments
4. Analysis of Alternatives
5. Customer relevance
6. Logical Arguments
7. Consistency (explanations)
8. Accuracy
9. Visualization

E. Blasch, J. Sung, T. Nguyen, "Multisource AI Scorecard Table for System Evaluation," *AAAI FSS-20: Artificial Intelligence in Government and Public Sector*, Washington, DC, USA, 2020. arXiv:2102.03985

Figure 9.2

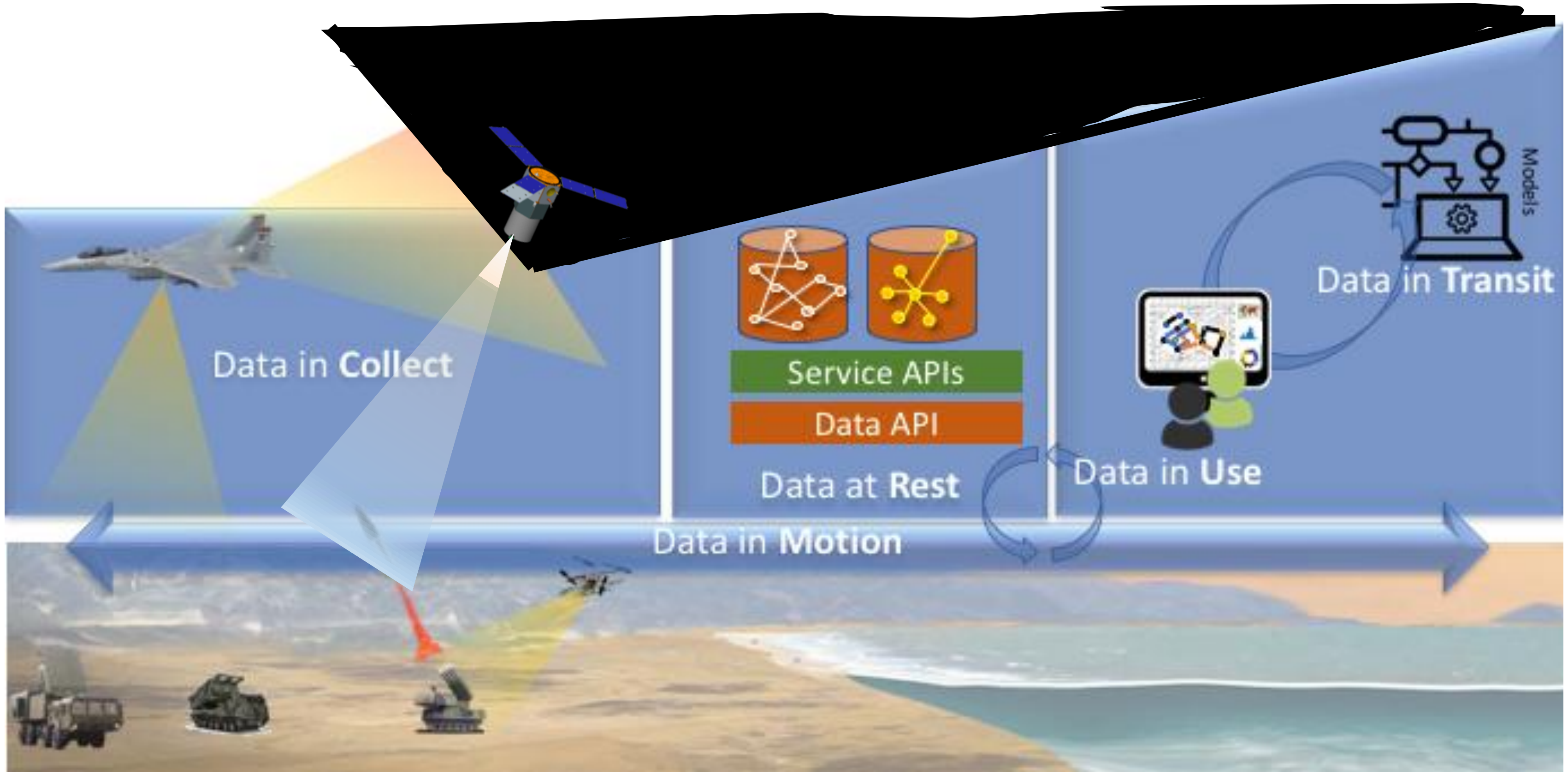


© Majumder, Blasch, Garren

Multi Domain Sensing

Figure 9.3

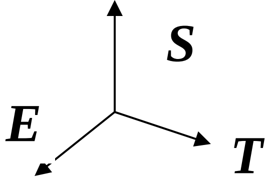
Evaluation for: Users, Missions, Sensor Control



Evaluation from: Operating conditions: Sensor, Environment, Target (SET)

B. Kahler, E. Blasch, and L. Goodwon, "Operating Condition modeling for ATR Fusion Assessment," *Proc. of SPIE*, 6571, 2007.

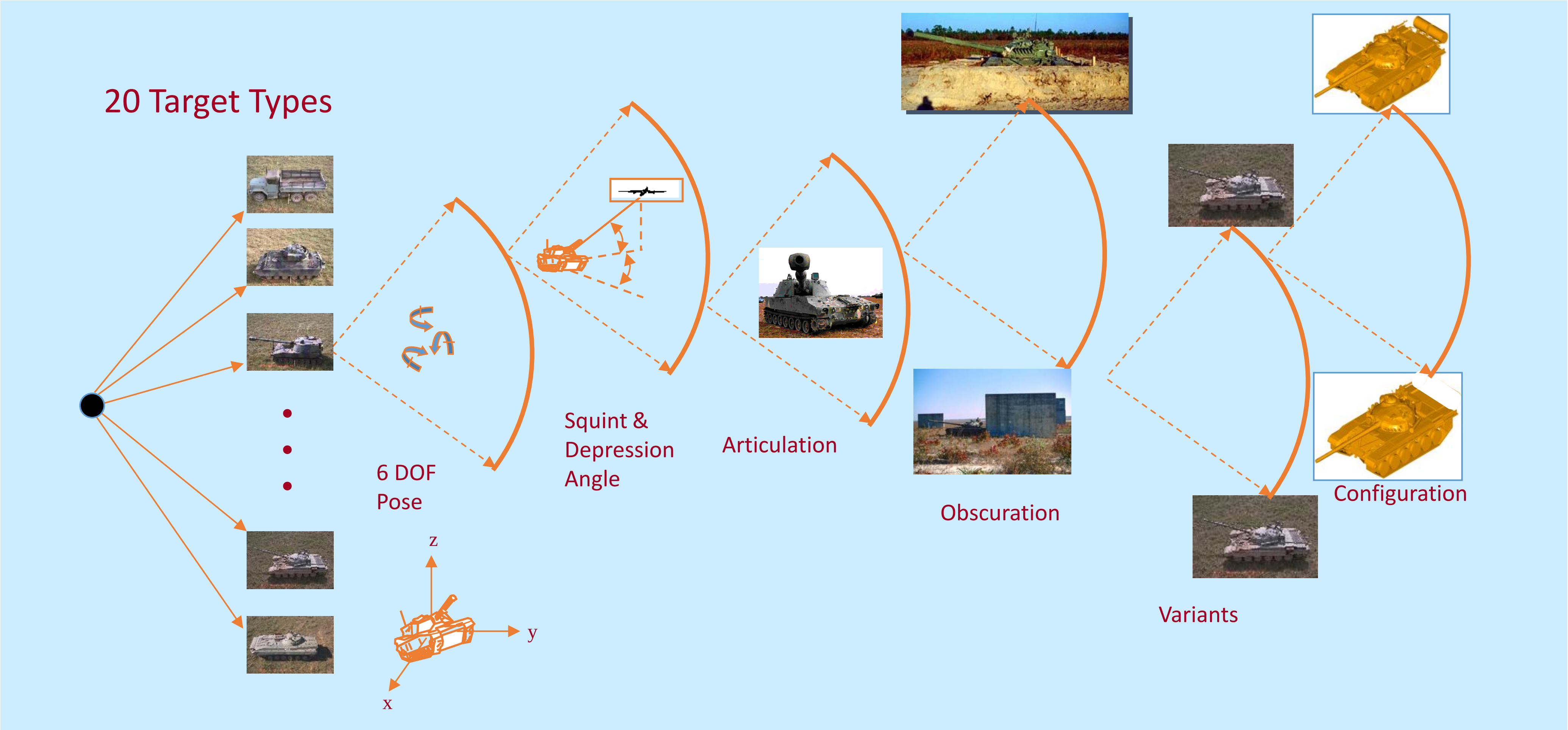
© Majumder, Blasch, Garren



Real world variability:

Extended Operating Conditions (EOC's)

Complexity – Operating Conditions (Constants, factors, MC)



$$N = 20 \times Q_1^6 \times Q_2^2 \times Q_3^2 \times Q_4^2 \times Q_5 \times Q_6 >> \text{Billions}$$

© Majumder, Blasch, Garren

Experimental Design

Complexity – Operating Conditions (Constants, factors, MC)

BASELINE (Constant)

Target

- Type of Target
- Target density

Environment

- Types of forest, Density
- Type of Terrain (slope)
- Season
- Weather
- Cultural vs. Natural

Sensor

- Acquisition Geometry
- Interference
- Image Quality
- # Images for baseline
- Time

FACTOR CHANGE

Target

- Change in Dynamics
- Change in Articulation

Environment

- Foliage Type
- Flat, Mountainous
- Change in Season
- Change in Weather
- Social / Polit. Context

Sensor

- Pointing accuracy
- Bias
- Resolution
- # of Looks
- Time Delay

Monte Carlo Variance

Target

- Dynamic Model Var.
- Articulation accuracy

Environment

- Foliage variance
- DTED variance
- Value of cloud
- Moisture content
- Planned AOI

Sensor

- GPS accuracy
- Noise
- Pixel variance
- # Look Variance
- Time Delay Variance

© Majumder, Blasch, Garren

Experimental Design

Complexity – Operating Conditions (Constants, factors, MC)

OC Category	Parameter	Flat Terrain	Urban Terrain
Targets	Targets	2, 10, 100	1, 2, 5
	Moving/Confusers	(1/0),(5/2), (100/1000)	(1/0), (2/10), (5/1000)
	Routes-Stop-Move	Variable	Variable
Sensors	Initial Start Points	Variable	Variable
	Bias	On/Off	On/Off
Environment	'No Fly' Zones	Variable Area Locations	Variable Area Locations

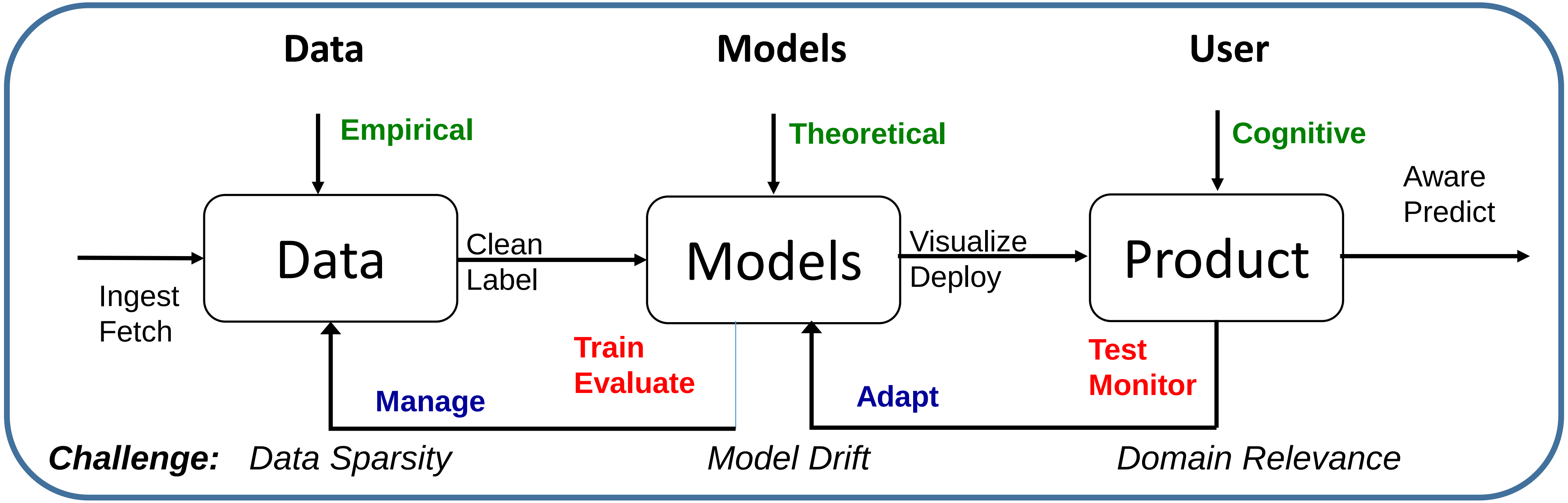
Design of Exp	Exp 1 (easy)	Exp 2 (not so easy)	Exp 3 (Difficult)		Exp N (Most Difficult)
Targets	5 known	5 /1 unknown	10 /4 unknown		K known/U unknown
Confusers	2 confusers	2 confusers 1 unknown	10 confusers 5 unknown		C confusers U unknowns, D decoys
PD/PID	95%/92%	90%/80%	70%/50%		100%/100%
Learning Rate	None	Limited data	Variable		Variable
Run Time	Unlimited	1 minute	30 seconds		milliseconds
Compute Energy	Unlimited	50 watts/2 lbs	10 Watts / 1 lb		mWatt/grams

© Majumder, Blasch, Garren

AI (ATR) Pipeline

Figure 9.4

- **Evaluation Analysis based on the Mission Awareness support**
  - **Agility –** e.g., Open Architectures, signal processing (Systems Software)
  - **Autonomy –** e.g., Context Awareness, health monitoring (Algorithms)
  - **Multi-domain –** e.g., Coordinated Sensing (Modeling, Instrumentation)



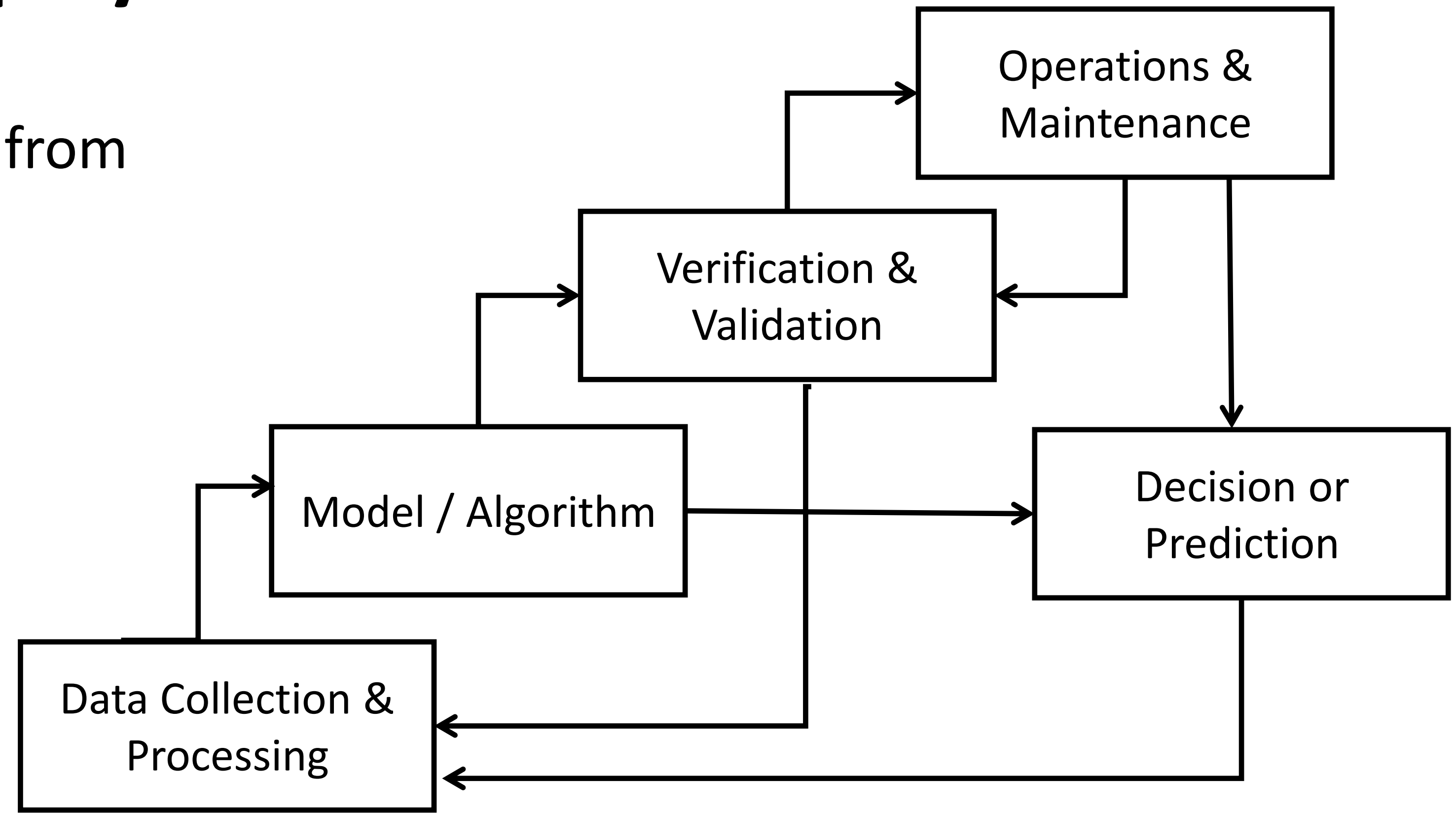
E. Blasch, J. Sung, T. Nguyen, C. P. Daniel, A. P. Mason, " Artificial Intelligence Strategies for National Security and Safety Standards," *AAAI Fall Symposium Series*, Nov. 2019. https://arxiv.org/ftp/arxiv/papers/1911/1911.05727.pdf

© Majumder, Blasch, Garren

Product Assessment and Deployment

Concerns with DL ATR deployment

Figure 9.5



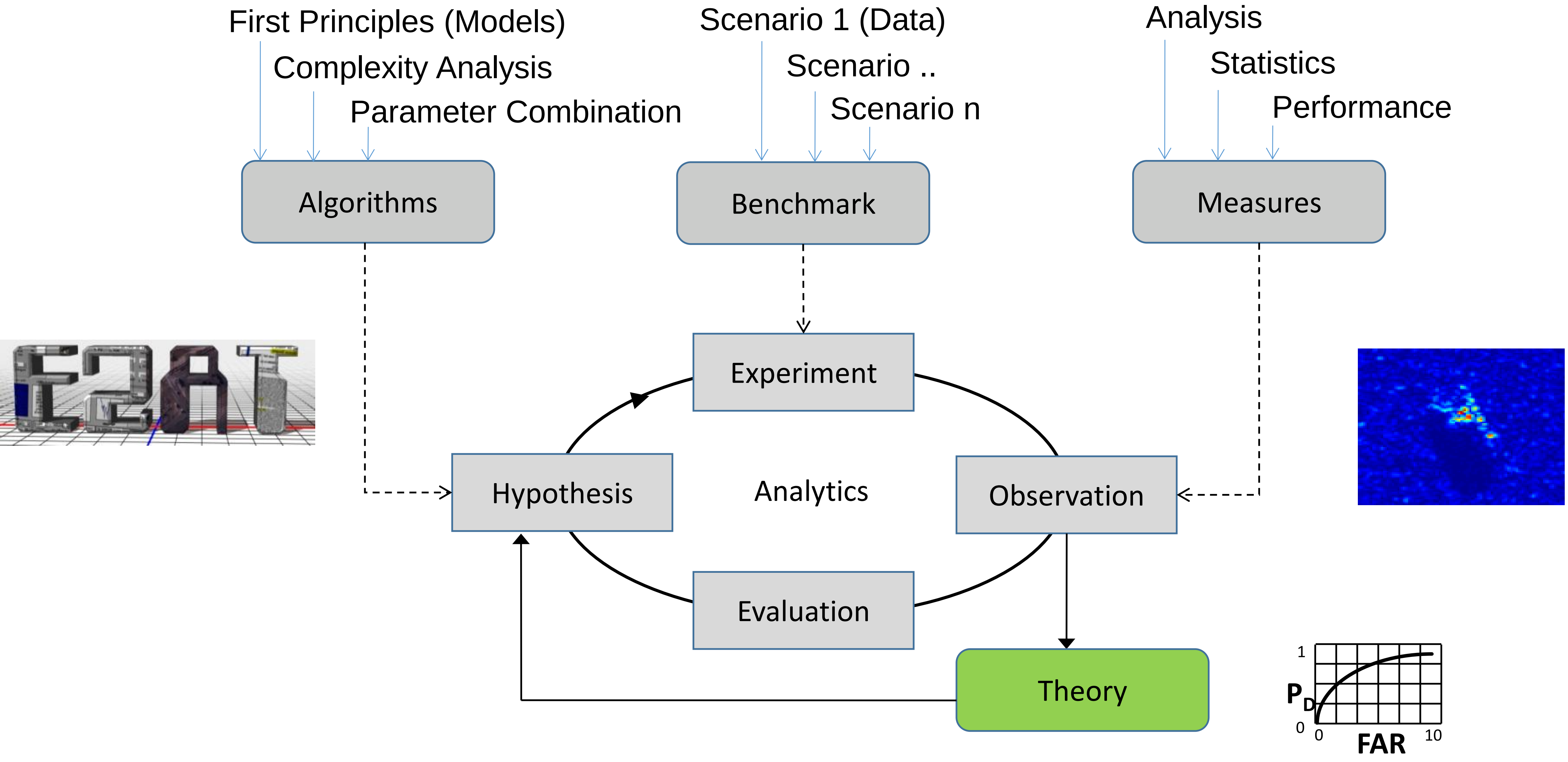
- **Data Collection - collect**
  - What data, how much, where from
- **Models - build**
  - What type, level of fidelity
- **V&V - measure**
  - How, When, conditions
- **O&M**
  - Who, cost, sustainment
- **Decision**
  - Intended use and support (Governance (CC), Design (Eng), Development (Soft), user)

Continual Life Cycle Design

© Majumder, Blasch, Garren

Testing and Evaluation

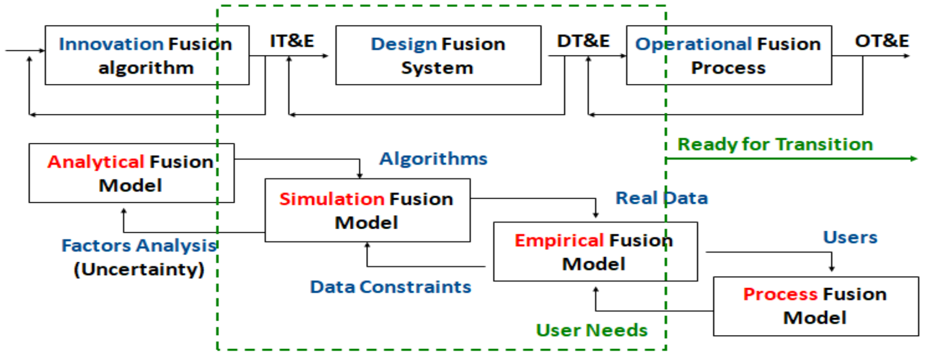
Figure 9.6



© Majumder, Blasch, Garren

Test and Evaluation Methods

Verification and Validation (V&V)



Goals (test and evaluation)

- Innovative test and evaluation – exploratory (IT&E)
- Design test and evaluation – development (**DT&E**)
- Operational test and evaluation – utilization (OT&E)
- Software test and evaluation – capability (ST&E)

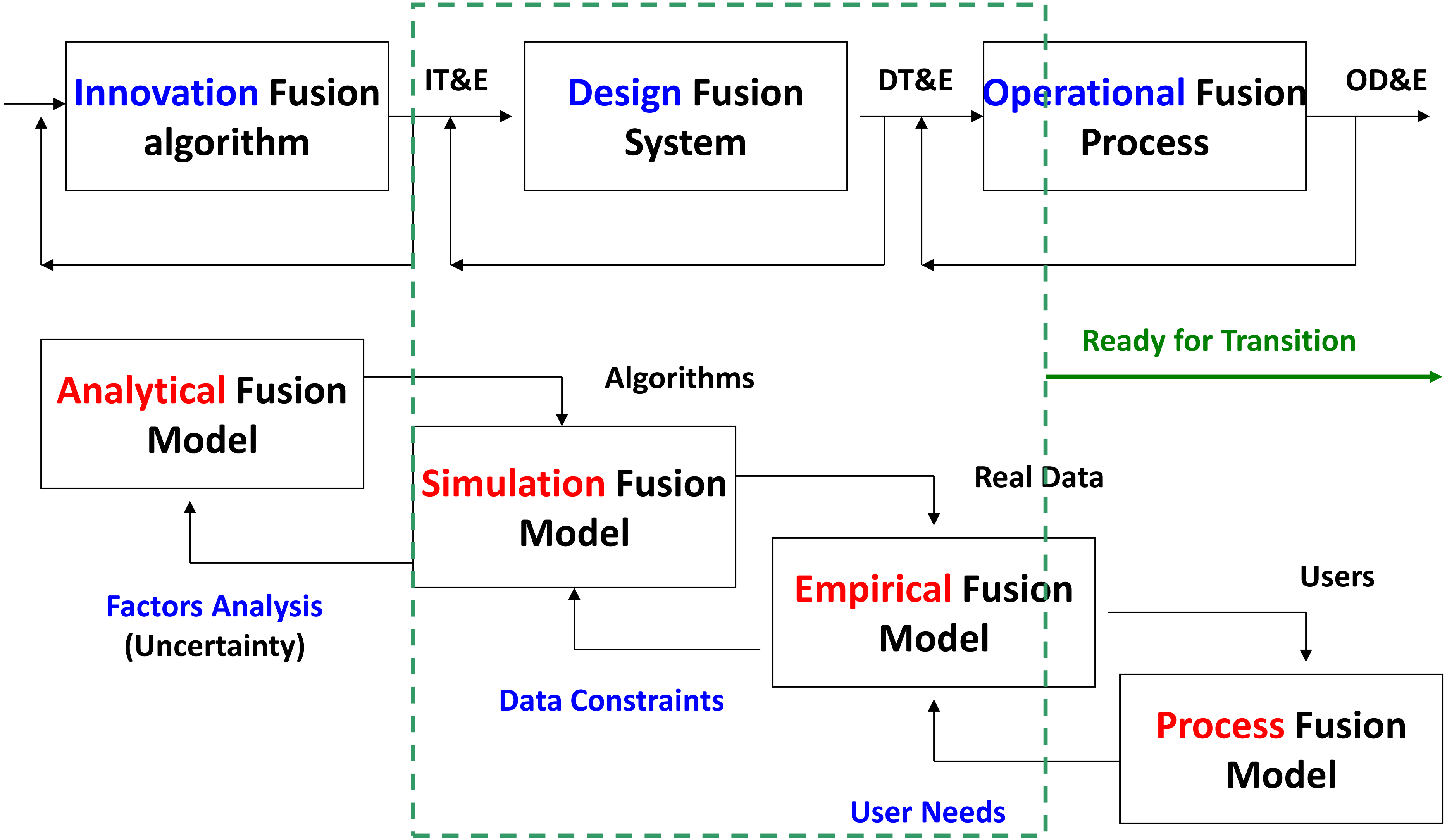
Data (model instantiation)

- Analytical model – mathematical performance analysis
- Simulation model – robust parameter exploration
- Empirical model – real data collected from sensors
- Process model – observations of humans w/ systems

© Majumder, Blasch, Garren

Design Test and Evaluation

Figure 9.7



© Majumder, Blasch, Garren

Uncertainty Representation and Reasoning Evaluation Framework (URREF) (ontology) Metrics

Figure 9.8

URREF
- Common definitions
- Uncertainty focus

Evaluation Concerns
- Data
- Reasoning
- Handling
- Reporting
- Repeatability
- Interpretability
- Explainability

Data Criterion

Credibility
- Objectivity
- Observational Sensitivity
- Self confidence

Quality
- Accuracy
- Precision
- Veracity

Relevance to Problem

Weight of Evidence

Reasoning Criterion

Performance
- Throughput
- Timeliness

Computational Cost

Correctness

Consistency

Scalability

Data Handling Criterion

Interpretation

Traceability

Representation Criterion

Expressiveness
- Assessment
- Outcomes
- Dependency
- Relational
- Higher-order Uncertainty
- Inclusiveness

Adaptability

Compatibility

Simplicity

Knowledge Handling

© Majumder, Blasch, Garren

Welcome to the ETURWG Home | ETURWG (gmu.edu)

Measures of Performance and Measures of Effectiveness

Measures of Performance : Machine

Measures of Effectiveness: User

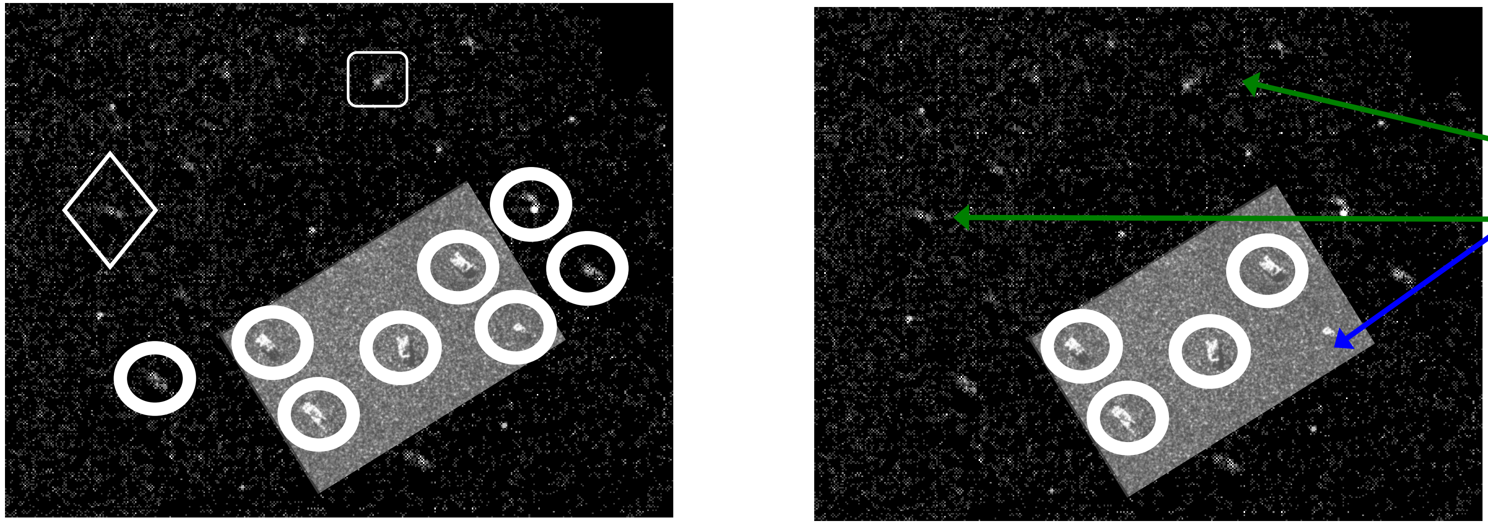
	Measures of Performance		Measures of Effectiveness		
	Credibility/Confidence	**Accuracy**	**Timeliness**	**Throughput**	**Cost**
Detection	*Probability of False Alarm (PFA)* False Alarm Rate (FAR)	*Probability of Detection (PD)*	*Trainability*	*Robustness*	*Database size*
Classification (Recognition)	PCC Conditioned on detection PCC Conditioned on target database	Probability of Correct Classification (PCC)	Time to acquisition	Complexity	Energy
Identification	Probability of correct association Fingerprinting	Probability of identification (PID)	Track lifetime	Security	Assurance
Tracking	Track Length Gap time	Track Purity	Revisit Rate	Bandwidth	Number of sensors

© Majumder, Blasch, Garren

Measure of Performance (Example)

Figure 9.9

A useful example from the MSTAR data.



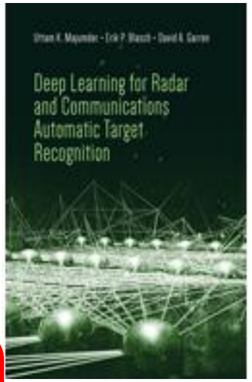
Machine = 80% performance

Human can't trust and has to check all the dots, thus it is 5/23 = 30% effective

- Assume there are ten targets. If there are 8 detections(◯), **then P_D = 8/10**, while for the 1 false alarm (□),P_{FA} = 1/10 and one [Missed Detections (◇)] , P_{FA} = 1/10.
- Using the extraction location of only the 5 targets and the ability to classify the target, then the conditional analysis of P(classification | detection) = 4/5.
- Once the detections are determined, then the target type classification are obtained via a confusion matrix.

© Majumder, Blasch, Garren

Chapter 9 Outline



© Majumder, Blasch, Garren

Confusion Matrix

Figure 9.10

Decisions								
Test Objects (truth)		Detection Accept						**Detection Reject**
		Target					Non-Target	
		Class Accept				Class Reject		
		Type 1	**Type 2**	**Type 3**	**OTHER**			
	Type 1	$\mathbf{M_{1,1}}$	$\mathbf{M_{1,2}}$	$\mathbf{M_{1,3}}$	$\mathbf{M_{1,O}}$	$\mathbf{M_{1,R\text{-}CL}}$	$\mathbf{M_{1,NT}}$	$M_{1,R\text{-}Det}$
Target	**Type 2**	$\mathbf{M_{2,1}}$	$\mathbf{M_{2,2}}$	$\mathbf{M_{2,3}}$	$\mathbf{M_{2,O}}$	$\mathbf{M_{2,R\text{-}CL}}$	$\mathbf{M_{2,NT}}$	$M_{2,R\text{-}Det}$
	Type 3	$\mathbf{M_{3,1}}$	$\mathbf{M_{3,2}}$	$\mathbf{M_{3,3}}$	$\mathbf{M_{3,O}}$	$\mathbf{M_{3,R\text{-}CL}}$	$\mathbf{M_{3,NT}}$	$M_{3,R\text{-}Det}$
	Other	$\mathbf{M_{O,1}}$	$\mathbf{M_{O,2}}$	$\mathbf{M_{O,3}}$	$\mathbf{M_{O,O}}$	$\mathbf{M_{O,R\text{-}CL}}$	$\mathbf{M_{O,NT}}$	$M_{O,R\text{-}Det}$
Non-Target	**Type 4**	$\mathbf{M_{4,1}}$	$\mathbf{M_{4,2}}$	$\mathbf{M_{4,3}}$	$\mathbf{M_{4,O}}$	$\mathbf{M_{4,R\text{-}CL}}$	$\mathbf{M_{4,NT}}$	$M_{4,R\text{-}Det}$
	Type 5	$\mathbf{M_{5,1}}$	$\mathbf{M_{5,2}}$	$\mathbf{M_{5,3}}$	$\mathbf{M_{5,O}}$	$\mathbf{M_{5,R\text{-}CL}}$	$\mathbf{M_{5,NT}}$	$M_{5,R\text{-}Det}$

Numerator	
Denominator	

probability likelihood of classification (P_{CL}):

$$P(c_i|t_j) = \frac{\sum_{i=[1,2,3,\ldots,O]} M_{ii}}{\sum_{i=[1,2,3,\ldots,O]}\left(\sum_{j=[1,2,3,\ldots,O]} M_{ij}\right)}$$

© Majumder, Blasch, Garren

Confusion Matrix Analysis

Figure 9.11

User and Machine Tell Different Stories

Decisions								
Test Objects		Detection Accept						Detection Reject
		Target					Non-Target	
		Class Accept				Class Reject		
		Type 1	**Type 2**	**Type 3**	**OTHER**			
	Type 1							
Target	**Type 2**							
	Type 3							
	Other							
Non-Target	**Type 4**							
	Type 5							

Classification by ATR

Classification by Operator

Calculating Metrics

$$P_{\text{Declaration}} = \frac{A}{A+B}$$

$$P_{\text{FalseAlarm}} = \frac{E}{E+D}$$

$$P_{\text{Correct Classification}} = \frac{m{\bullet}P_{\text{D}}}{(m{\bullet}P_{\text{D}}) + P_{\text{FA}}}$$

$$CM = \begin{array}{c} \\ T1 \\ T2 \\ \\ TN \end{array} \begin{array}{ccccccc} T1 & T2 & & TN & TO & TR & Unk \\ A & B & \dots & B & B & B & U \\ E & D & \dots & D & D & D & U \\ \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots \\ E & D & D & D & D & D & U \end{array}$$

© Majumder, Blasch, Garren

Confusion Matrix Example

Figure 9.12

			Target					(Test)
			Class Accept				Class Reject	Unknown
(Truth)			**Type 1**	**Type 2**	**Type 3**	**OTHER**		
	Type 1	Friend	**0.80**	**0.03**	**0.01**	**0.01**	**0.02**	0.13
Target	**Type 2**	Friend	**0.01**	**0.85**	0.01	0.03	0.01	0.09
	Type 3	Neutral	**0.01**	0.01	**0.82**	0.01	0.03	0.12
	Other	Neutral	**0.03**	0.01	0.01	**0.90**	0.03	0.02
Non-Target	**Type 4**	Foe	**0.01**	0.02	0.04	0.21	0.30	0.42
	Type 5	Foe	**0.20**	0.20	0.20	0.02	0.14	0.24

P (Declaration) = A / (A + B) = 0.08 / [0.8 + (0.03+0.01+0.01+0.02))] = 0.92

P (False Alarm) = E / (E + D) **=** = 0.26 / [0.26+4.65] = 0.053

E = 0.01+0.01+0.03+0.01+0.20 = 0.26

U is the entire right column of 0.13+0.09+0.12+0.02+0.42+0.24=**1.02**.
Finally, *D* is the remaining value *D* = 6 – (0.07) – (0.26) – (1.02) = **4.65**.

© Majumder, Blasch, Garren

Confusion Matrix Example

Calculating Metrics

$$P_{\text{Declaration}} = \frac{A}{A+B}$$

$$P_{\text{FalseAlarm}} = \frac{E}{E+D}$$

$$P_{\text{Correct Classification}} = \frac{m \bullet P_{\text{D}}}{(m \bullet P_{\text{D}}) + P_{\text{FA}}}$$

$$\begin{array}{cc} & \begin{array}{ccccccc} T1 & T2 & & TN & TO & TR & Unk \end{array} \\ \mathrm{CM} = \begin{array}{c} T1 \\ T2 \\ \\ TN \end{array} & \begin{array}{|ccccccc|} \hline A & B & \cdots & B & B & B & U \\ E & D & \cdots & D & D & D & U \\ \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots \\ E & D & D & D & D & D & U \\ \hline \end{array} \end{array}$$

$$\begin{array}{cc} & \begin{array}{ccccccc} T1 & T2 & & TN & TO & TR & Unk \end{array} \\ \mathrm{CM} = \begin{array}{c} T1 \\ T2 \\ \\ TN \end{array} & \begin{array}{|ccccccc|} \hline D & E & \cdots & D & D & D & U \\ B & A & \cdots & B & B & B & U \\ \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots \\ D & E & D & D & D & D & U \\ \hline \end{array} \end{array}$$

Target 1: **0.80**.

***B* = 0.03+0.01+0.01+0.02** = **0.07**.

***E* = 0.01+0.01+0.01+0.03+0.01+0.20** =**0.26**.

U is the entire right column of 013+0.09+0.12+0.02+0.42+0.24=**1.02**.

D is the remaining value *D* = **4.65**.

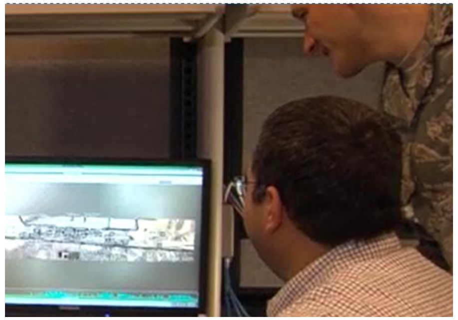
Target 1: P_{Dec} = 0.8/ [0.8+0.07] = **0.92.** P_{FA} = 0.26 / [0.26+4.65] = 0.053.

Let $m = 1$, then P_{CC} = 0.92/[0.92 + 0.05] = 0.95.

Target 2 : P_{DEC} = 0.**93**, P_{FA} = 0.055, and P_{CC} = 0.94.

© Majumder, Blasch, Garren

Confusion Matrix Example



Calculating Metrics: column analysis can utilize the a priori probabilities

$$P_{ID}(c_j|t_i) = \frac{P(type_i|classification_j)\,P(type_i)}{\sum_{i=\{1,2,3,\ldots,n\}} P(type_i|classification_j)P(type_i)}$$

For example, let P_F be the prior probability of Friendly targets (1, 2), P_N be the prior probability of Neutral targets (3, 4) and P_H is the prior probability of Hostile (Foe) targets (5, 6).

For a **non-threat environment** P_F=4P_H, (0.4, 0.4, 0.0, 0.0, 0.1 and 0.1), so that:

$$P_{ID}(c_i|t_j) = \frac{(0.8)\,(0.4)}{(0.8+0.01)(0.4)+(0.01+0.03)(0.0)+(0.01+0.20)(0.1)} = \frac{0.32}{0.09+0.0602} = \boxed{93\%}$$

With a **threat environment of P_H= 4P_F,** the priors for each true target types are (0.1, 0.1,0.00, 0.00, 0.40 and 0.40). Then, for target type 1:

$$P_{ID}(c_i|t_j) = \frac{(0.8)\,(0.1)}{(0.8+0.01)(0.1)+(0.01+0.03)(0.0)+(0.01+0.20)(0.4)} = \frac{0.08}{0.081+0.084} = \boxed{48\%}$$

Hence, operators acting on an ATR decision system should realize that, **without using a prior knowledge**, the results can be significantly inaccurate (e.g., 48% versus 93%). Hence, the ATR analysis from the machine presents both 93% for classification and 93% for known status-quo engagement. However, with the unknown, the user **should be aware that the ATR results are 75% credible**, and then as a threat (e.g., hostile) situation unfolds, the ATR results would be only 50% credible.

© Majumder, Blasch, Garren

Confusion Matrix Analysis

Figure 9.13

Decisions								
Test Objects		Detection Accept						Detection Reject
		Target					Non-Target	
		Class Accept				Class Reject		
		Type 1	**Type 2**	**Type 3**	**OTHER**			
	Type 1	**0.80**						
Target	**Type 2**		**0.85**					
	Type 3			**0.82**				
	Other				**0.90**			
Non-Target	**Type 4**							
	Type 5							

Classification by ATR

0.92
0.93
0.93
0.92

0.75 0.76 0.75 0.76

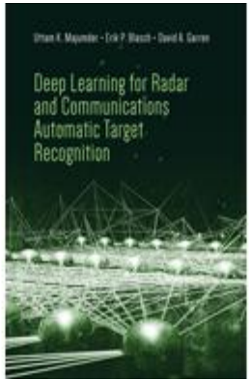
Classification by Operator

Type 1 classification ROW (for the ATR) is 0.8/(0.8+0.03+0.01+0.01+0.02) = 92%; with an uncertainty of 15%.
Type 1 classification: COLUMN (for the user) is 0.8/(0.8+0.01+0.01+0.03+0.01+0.02) = 75%.

ATR provides overconfidence with an average of 92%;
while the **user results are conservative around 75%**

© Majumder, Blasch, Garren

Chapter 9 Outline



© Majumder, Blasch, Garren

Receiver Characteristic Operating Curve

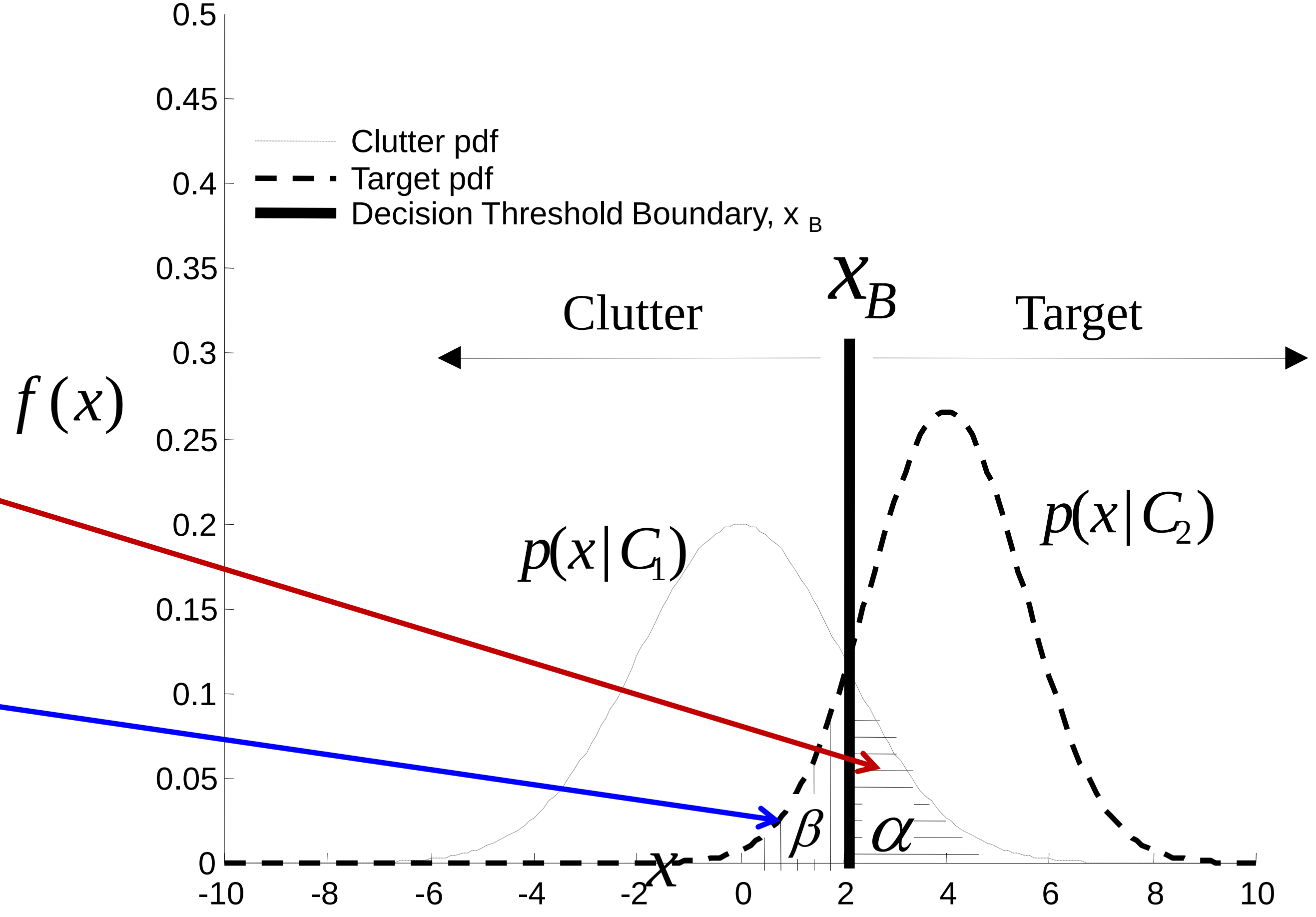
Figure 9.14

Signal Detection Theory:

The **clutter** pdf $p(x|C_1)$ is the conditional pdf representing the distribution of clutter objects (class 1), while the **target** pdf$p(x|C_2)$ is the conditional pdf representing the distribution of target objects (class 2).

Type I Error (α): Misclassifying an actual **clutter object** as a target (False Positive, FP, or False Alarm, FA)

Type II Error (β): Misclassifying an actual **target object** as clutter (False Negative, FN)

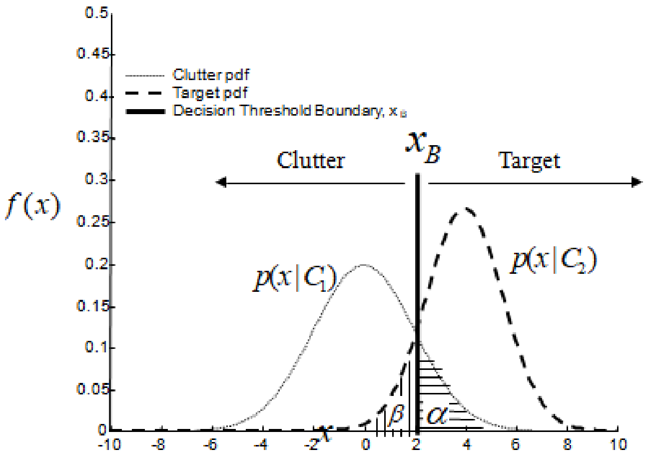


© Majumder, Blasch, Garren

Receiver Characteristic Operating Curve

Metrics

Figure 9.14



		Classified as (reported)	
		Clutter (C) Normal (N) Null Hypothesis H_0	**Target (T)** Abnormal (A) Alt Hypothesis H_1
Actual (Truth)	**Clutter (c)** Normal (n) Null Hypothesis H_0	True Negative (P_{TN}) Correct Rejection (P_{CR}) Specificity Confidence ($1-\alpha$) P(C|c) P(N|n)	False Positive (P_{FP}) **False Alarm (P_{FA})** 1 - Specificity Level of Significance (α) P(T|c) P(A|n)
	Target (t) Abnormal (a) Alt Hypothesis H_1	False Negative (P_{FN}) Missed Detection (P_M) 1 - Sensitivity 1 - power (β) P(C|t) P(N|a)	True Positive (P_{TP}) **Detection (P_D)** Sensitivity Power ($1 - \beta$) P(T|t) P(A|a)

© Majumder, Blasch, Garren

Receiver Characteristic Operating Curve (Example)

Metrics

			Target					(Test)
			Class Accept				Class Reject	Unknown
(Truth)			Type 1	Type 2	Type 3	OTHER		
	Type 1	Friend	**0.80**	0.03	0.01	0.01	0.02	0.13
Target	**Type 2**	Friend	0.01	**0.85**	0.01	0.03	0.01	0.09
	Type 3	Neutral	0.01	0.01	**0.82**	0.01	0.03	0.12
	Other	Neutral	0.03	0.01	0.01	**0.90**	0.03	0.02
Non-Target	**Type 4**	Foe	0.01	0.02	0.04	0.21	0.30	0.42
	Type 5	Foe	0.20	0.20	0.20	0.02	0.14	0.24

U

Using the CM then the detection comes from the correct results,
while the false alarms are the errors from the columns (T):

U is the entire right column of 0.13+0.09+0.12+0.02+0.42+0.24=**1.02**.

Denominator = Total - U = 6 – 1.02 = 4.98

False Alarm = P(T|c) = 0.26 +0.27+0.27+0.73 / (4.98) = 0.13

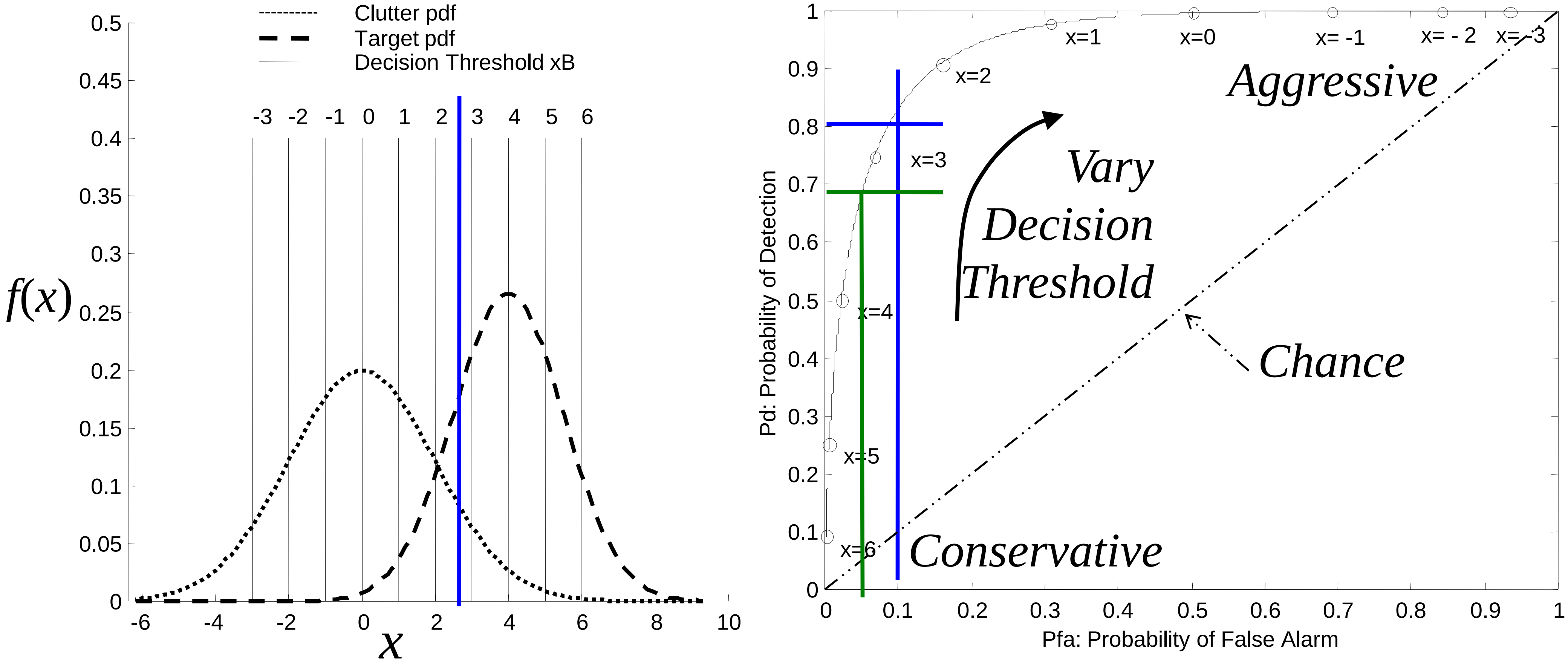
Detection = P(T|t) = 0.8+0.85+0.82+0.90+(0.44)/ (4.98) = 0.77

© Majumder, Blasch, Garren

ROC Generation

Figure 9.15

Generation of ROC Curve (right) using two 1-D normal distributions (left). Clutter: N(2,0); Target: N(4,1.5)



NOTE: If a classifier is listed with a detection of P_D= 80%, what does this mean? (P_D of 80% is at P_{FA} of 10%)

OPERATOR: wants a 5% P_{FA}, then the P_D is actually somewhere near 65%.

Thus, the ATR designer would want the operator to believe it is PD = 80%, but it is actually at 65%

© Majumder, Blasch, Garren

3D ROC

After setting performance measures for Pd, Pfa, the additional parameters *z* are added:

$$g = g(z) = \left\{ \left(\hat{P}_{FA}(z), \hat{P}_D(z), \hat{P}_3(z) \right) \middle| z \in Z \right\}$$

Additional parameters *z* examples (**time, space, operating condition**)

With the sum of probabilities: define outputs o_1 and o_2 for each exemplar ($o_1 + o_2 = 1$), and define thresholds:

$$class = \begin{cases} class\ 2 \ \text{ if } \ o_2 > t_u \\ class\ 1 \ \text{ if } \ o_2 < t_l \\ reject \ \text{ if } \ t_l \le o_2 \le t_u \end{cases} \qquad t_l = t_c - \frac{1}{2} z \qquad t_u = t_c + \frac{1}{2} z$$

To generate the 3-D ROC trajectory, set $t_c = 0.5$ and vary from 0 (no rejections) to 1 (all exemplars rejected).
Probability of Rejection (P_r) the probability that the image is rejected as unknown or is too difficult to classify

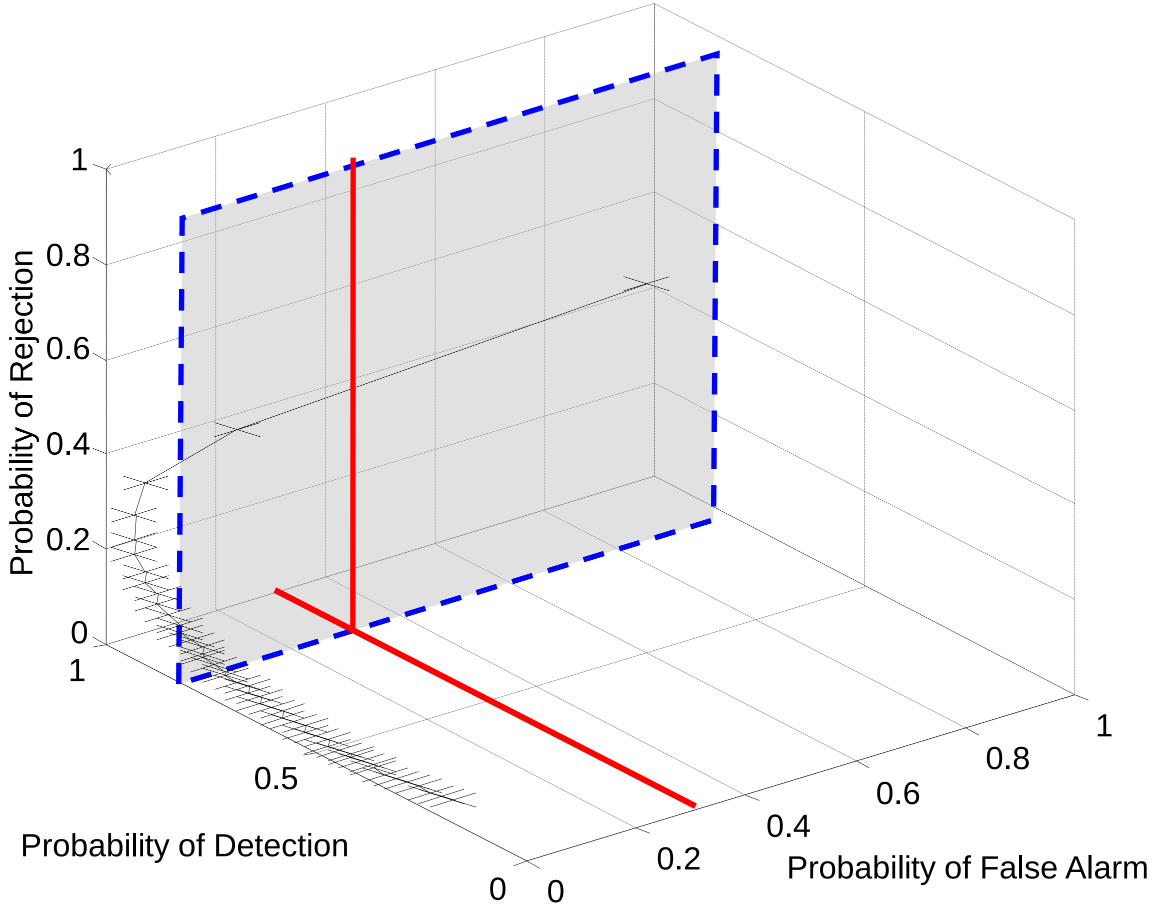
E. P. Blasch, S. Alsing, and R. Bauer, “Comparison of bootstrap and prior probability synthetic data balancing method for SAR target recognition,” *Proc. SPIE, Vol, 3721,* April 1999.

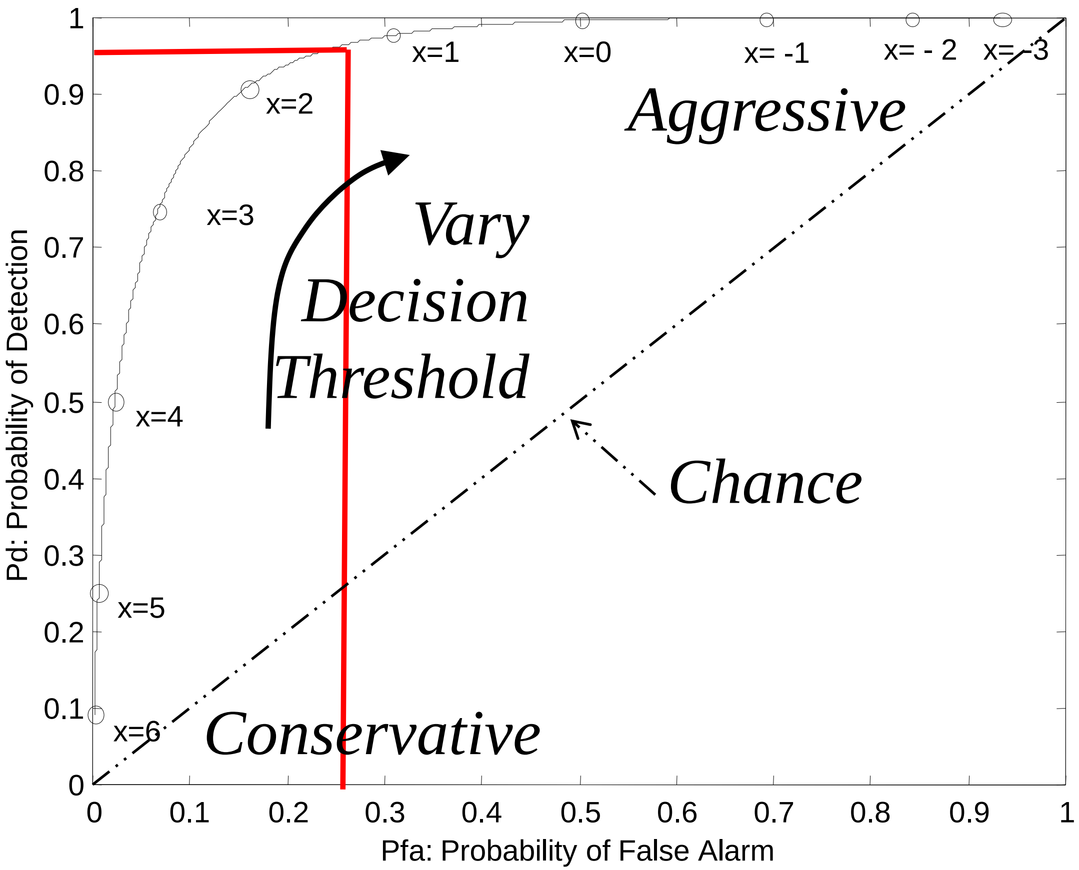
S. Alsing, E. P. Blasch, R. Bauer, “Three-Dimensional Receiver Operating Characteristic (ROC) trajectory concepts for the Evaluation of Target Recognition algorithms faced with the Unknown target detection problem,” *Proc. SPIE* 3718, 1999

© Majumder, Blasch, Garren

3D ROC

Figure 9.16





E. P. Blasch, S. Alsing, and R. Bauer, "Comparison of bootstrap and prior probability synthetic data balancing method for SAR target recognition," *Proc. SPIE, Vol, 3721,* April 1999.

S. Alsing, E. P. Blasch, and R. Bauer, "Three-Dimensional Receiver Operating Characteristic (ROC) trajectory concepts for the Evaluation of Target Recognition algorithms faced with the Unknown target detection problem," *Proc. SPIE,* Vol. 3718, 1999

© Majumder, Blasch, Garren

z

Precision-Recall Matrix

OTHER ROC metrics:

Area under the curve (AUC)

The **F-measure** is determined from a precision-recall (curve) using the confusion matrix and the number of tests

- True positive (TP), true negative (TN), false positive (FP), and false negative (FN).
- The metrics for precision and recall are determined from the CM as a comprehensive F-metric.

		Classified as (reported)	
		Clutter (C) Normal (N) Null Hypothesis H_0	Target (T) Abnormal (A) Alt Hypothesis H_1
Actual (Truth)	Clutter (c) Normal (n) Null Hypothesis H_0	True Negative (P_{TN}) Correct Rejection (P_{CR}) P(C|c)	False Positive (P_{FP}) **False Alarm (P_{FA})** P(T|c)
	Target (t) Abnormal (a) Alt Hypothesis H_1	False Negative (P_{FN}) Missed Detection (P_M) P(C|t)	True Positive (P_{TP}) **Detection (P_D)** P(T|t)

© Majumder, Blasch, Garren

z

Precision-Recall Matrix (Example)

Figure 9.17

ROC, which include the area under the curve (AUC) and the F-metric. The F-measure is determined from a precision-recall (curve) using the confusion matrix and the number of tests

			Target					(Test)
			Class Accept				Class Reject	Unknown
(Truth)			**Type 1**	**Type 2**	**Type 3**	**OTHER**		
	Type 1	Friend	**0.80 = TP**	0.03 = FP	0.01 = FP	0.01 = FP	0.02 = TN	0.13 = FN/TN
Target	**Type 2**	Friend	0.01 = FP	**0.85 = TP**	0.01 = FP	0.03 = FP	0.01 = TN	0.09 = FN/TN
	Type 3	Neutral	0.01 = FP	0.01 = FP	**0.82 = TP**	0.01= FP	0.03 = TN	0.12 = FN/TN
	Other	Neutral	0.03 = FP	0.01 = FP	0.01 = FP	**0.90 = TP**	0.03 = TN	0.02 = FN/TN
Non-Target	**Type 4**	Foe	0.01 = FP	0.02 = FP	0.04 = FP	0.21 = FP	0.30 = TN	0.42 = FN/TN
	Type 5	Foe	0.20 = FP	0.20 = FP	0.20 = FP	0.02 = FP	0.14 = TN	0.24 = FN/TN

© Majumder, Blasch, Garren

Precision-Recall Matrix (Example)

Figure 9.17

ATR Metrics

			Target					(Test)
			Class Accept				Class Reject	Unknown
(Truth)			**Type 1**	**Type 2**	**Type 3**	**OTHER**		
	Type 1	Friend	**0.80 = TP**	0.03 = FP	0.01 = FP	0.01 = FP	0.02 = TN	0.13 = FN/TN
Target	**Type 2**	Friend	0.01 = FP	**0.85 = TP**	0.01 = FP	0.03 = FP	0.01 = TN	0.09 = FN/TN
	Type 3	Neutral	0.01 = FP	0.01 = FP	**0.82 = TP**	0.01= FP	0.03 = TN	0.12 = FN/TN
	Other	Neutral	0.03 = FP	0.01 = FP	0.01 = FP	**0.90 = TP**	0.03 = TN	0.02 = FN/TN
Non-Target	**Type 4**	Foe	0.01 = FP	0.02 = FP	0.04 = FP	0.21 = FP	0.30 = TN	0.42 = FN/TN
	Type 5	Foe	0.20 = FP	0.20 = FP	0.20 = FP	0.02 = FP	0.14 = TN	0.24 = FN/TN

	Yes	No
Truth(Yes)	TP = 337	FN = 51
Truth(No)	FP = 108	TN = 104

$$Precision(P) = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}} = \frac{337}{337 + 108} = 76\%$$

$$Recall(R) = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} = \frac{337}{337 + 51} = 87\%$$

$$Error(E) = \frac{\mathrm{FP} + \mathrm{FN}}{\mathrm{n}} = \frac{108 + 51}{600} = 27\%$$

$$Accuracy(A) = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{n}} = \frac{337 + 104}{600} = 74\%$$

the F-measure or balanced F-score (**F_1 score**),which is the harmonic mean of precision and recall, where an F_1 score reaches its best value at 1 (perfect precision and recall) and worst at 0.

- ***Precision*** is a measure of **how many selected items are relevant**,
- ***Recall*** is how many **relevant items are selected**.

The general formula for positive real β modifies the weights of precision and recall:

$$F_\beta = (1 + \beta^2)\frac{\mathrm{Precision} * \mathrm{Recall}}{(\beta^2 * \mathrm{Precision}) + \mathrm{Recall}}$$

- $F_{0.5}$ weighs recall lower than precision (by attenuating the influence of false negatives) and
- F_2weighs recall higher than precision (by placing more emphasis on false negatives).
- Hence, $F_{0.5}$= **0.78**, F_1= 0.81, and F_2= 0.85.

© Majumder, Blasch, Garren

Confusion Matrix Fusion

- **Enhance metrics: (data fusion)**
  - Multiple looks **at** the target (getting closer)
  - Multiple looks **around** the target (circle)
  - Multiple looks **of** the target (dynamic tracking)
  - Need to fuse the confusion matrices for each iterative look (Bayesian update)

Improve precision, recall, etc.

CM 15 PCNN

	TARGET1	TARGET2	TARGET3	TARGET4	TARGET5	TARGET6	TARGET7	TARGET8	TARGET9	TARGET10	TARGET11	OTHER
TARGET 1	0.69	0.1	0.05	0.02	0.02	0.02	0.02	0.04	0.02	0	0.02	0
TARGET 2	0.07	0.66	0.12	0.01	0.02	0.03	0	0.02	0.04	0	0.03	0.01
TARGET 3	0.04	0.13	0.63	0.03	0	0	0	0.02	0.06	0.01	0.04	0.03
TARGET 4	0.02	0.01	0.03	0.65	0	0	0	0	0.08	0.02	0.16	0.03
TARGET 5	0.02	0.04	0	0	0.73	0.16	0.05	0	0	0	0	0
TARGET 6	0.02	0.04	0	0	0.17	0.62	0.14	0	0	0	0	0
TARGET 7	0.04	0	0	0	0.1	0.26	0.6	0	0	-0	0	0
TARGET 8	0.05	0.04	0.05	0	0	0	0	0.82	0.03	0	0	0.01
TARGET 9	0.01	0.05	0.07	0.09	0	0	0	0.02	0.7	0	0.05	0.01
TARGET 10	0	0	0.04	0.04	0	0	-0	0	0	0.54	0.19	0.19
TARGET 11	0.01	0.02	0.02	0.08	0	0	0	0	0.02	0.04	0.75	0.05
NOT-IN-LIB	0	0.01	0.04	0.03	0	0	0	0.01	0.01	0.09	0.11	0.7

CM 17 PCNN

	TARGET1	TARGET2	TARGET3	TARGET4	TARGET5	TARGET6	TARGET7	TARGET8	TARGET9	TARGET10	TARGET11	OTHER
TARGET 1	0.76	0.11	0.03	0.01	0.01	0.04	0.01	0.01	0	0.02	0.01	0
TARGET 2	0.21	0.75	0.02	0	0	0	0.01	-0	0	0.01	0	0
TARGET 3	0.05	0.02	0.86	0	0	0.06	0	0	0	0.01	0	0
TARGET 4	0.01	0	0	0.72	0.05	0.01	0.09	0	0	0.03	0.03	0.05
TARGET 5	0.01	0	0	0.02	0.77	0.1	0.04	0.01	0.03	0.01	0	0
TARGET 6	0.03	0	0.02	0.01	0.1	0.73	0.02	0.01	0.02	0.05	0	0
TARGET 7	0.01	0.01	0	0.05	0.05	0.03	0.83	0	0	0.01	0.01	0.01
TARGET 8	0.01	-0	0	0	0.03	0.03	0.01	0.91	0	0	0	0.02
TARGET 9	0	0	0	0	0.05	0.04	0	0	0.79	0.1	0	0.02
TARGET 10	0.02	0.01	0.01	0.02	0.02	0.06	0.01	0	0.07	0.74	0.05	0.01
TARGET 11	0.01	0	0	0.02	0	0	0.02	0	0	0.08	0.84	0.03
NOT-IN-LIB	0	0	0	0.05	0.01	0	0.02	0.02	0.03	0.01	0.03	0.83

CM CM PCNN

	TARGET1	TARGET2	TARGET3	TARGET4	TARGET5	TARGET6	TARGET7	TARGET8	TARGET9	TARGET10	TARGET11	OTHER
TARGET 1	0.79	0.11	0.03	0.01	0.01	0.02	0.01	0.01	0	0	0	0
TARGET 2	0.12	0.85	0.02	0	0	0.01	0	0	0	0	0	0
TARGET 3	0.04	0.03	0.9	0	0	0.01	0	0	0	0.01	0	0
TARGET 4	0.01	0	0	0.85	0	0	0	0	0.01	0.03	0.05	0.04
TARGET 5	0.01	0	0	0	0.82	0.12	0.05	0	0	0	0	0
TARGET 6	0.03	0	0	0	0.11	0.81	0.03	0	0	0.01	0	0
TARGET 7	0.01	0	0	0	0.05	0.04	0.89	0	0	0	0	0
TARGET 8	0.01	0	0	0	0	0	0	0.97	0	0	0	0.01
TARGET 9	0	0	0	0.01	0	0	0	0	0.94	0.02	0	0.01
TARGET 10	0.01	0	0.01	0.02	0	0	0	0	0.02	0.87	0.04	0.02
TARGET 11	0.01	0	0	0.02	0	0	0	0	0	0.08	0.85	0.03
NOT-IN-LIB	0	0	0	0.04	0	0	0	0.01	0.01	0.01	0.03	0.88

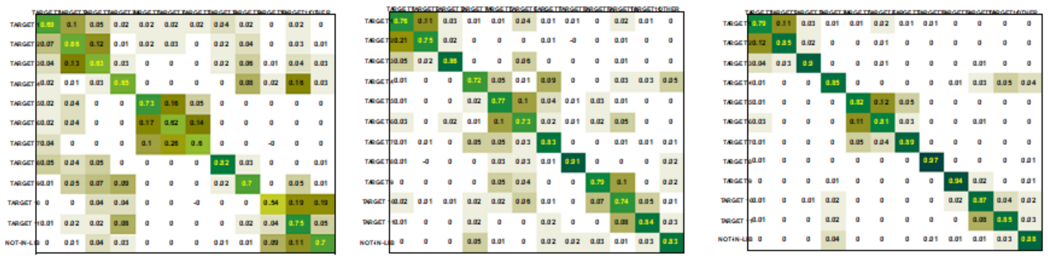
***see Chapter 4**

© Majumder, Blasch, Garren

Confusion Matrix Fusion

- **Fusion of Two Matrices**
- The priors and likelihoods are denoted as column vectors

$$p(\bar{o}) = \begin{bmatrix} p(o_1) \\ p(o_2) \\ \vdots \\ p(o_N) \end{bmatrix}; \quad p(z_j|\bar{o}) = \begin{bmatrix} p(z_j|o_1) \\ p(z_j|o_2) \\ \vdots \\ p(z_j|o_N) \end{bmatrix}$$



- Decisions

$$C = \begin{bmatrix} p(z_1|o_1) & p(z_2|o_1) & .. & p(z_M|o_1) \\ p(z_1|o_2) & p(z_2|o_2) & .. & p(z_M|o_2) \\ \cdots & \cdots & \ddots & \cdots \\ p(z_1|o_N) & p(z_2|o_N) & .. & p(z_M|o_N) \end{bmatrix}$$

- Joint likelihoods are similar column vectors, assuming independence for two confusion matrices *A* and *B*

$$p(z_j^A, z_k^B|\bar{o}) = \begin{bmatrix} p(z_j^A|o_1) \bullet p(z_k^B|o_1) \\ p(z_j^A|o_2) \bullet p(z_k^B|o_2) \\ \cdots \\ p(z_j^A|o_N) \bullet p(z_k^B|o_N) \end{bmatrix}$$

Algorithm 1: Confusion Matrix Fusion

```
function [d, pObarZaZb]=fuseCMdecisions(za, zb, Obar)
        CA = getConfusionMatrix(1);
        CB = getConfusionMatrix(2);
        pZaObar = CA(:,za);
        pZbObar = CB(:,zb);
        pZaZbObar = pZaObar .* pZbObar;
        posteriorNum = pZaZbObar .* pObar;
        posteriorDen = sum(posteriorNum);
        pObarZaZb = posteriorNum / posteriorDen;
        [junk, d] = max(pObarZaZb);
        Return
```

- Calculate *a posteriori* from with decision from max likelihood estimate

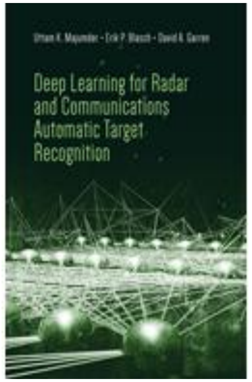
$$d_i = \arg\max_{j,k} p(o_i|z_j^A, z_k^B) \qquad p(\bar{o}|z_j^A, z_k^B) = \frac{p(z_j^A, z_k^B|\bar{o}_j)p(\bar{o}_j)}{\sum_{i=1}^{n} p(z_j^A, z_k^B|\bar{o}_i)p(\bar{o}_i)}$$

***code available**

B. Kahler and E. Blasch, "Decision-Level Fusion Performance Improvement from Enhanced HRR Radar Clutter Suppression," *J. of. Advances in Information Fusion*, Vol. 6, No. 2, pp. 101-118, Dec. 2011.

© Majumder, Blasch, Garren

Chapter 9 Outline

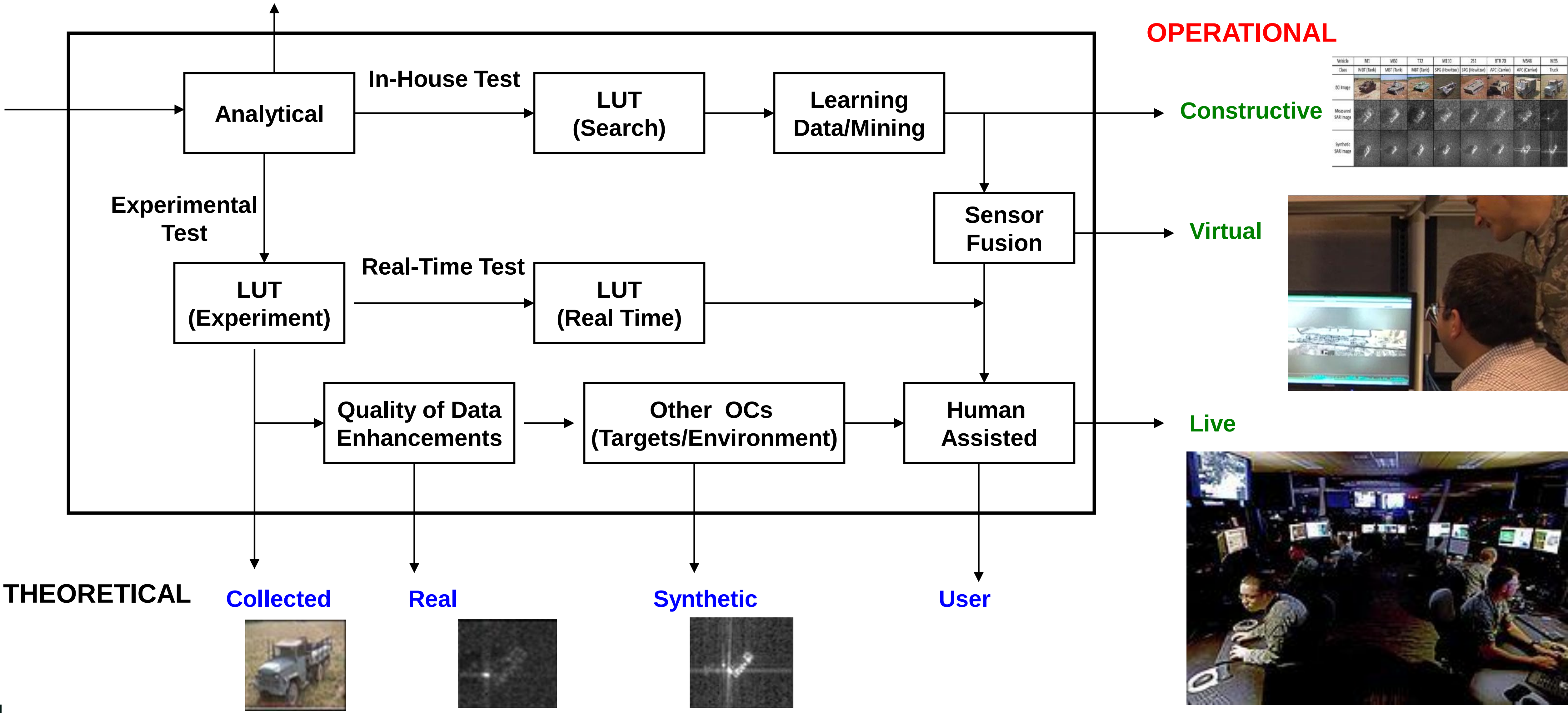


© Majumder, Blasch, Garren

Evaluation With Users

- **Need to assess analytical with results**

Figure 9.18



© Majumder, Blasch, Garren

National Interpretability Imagery Rating Scale (NIIRS)

National imagery interpretability rating scale (NIIRS) to assess the **quality of SAR data**. Given a SAR image with a NIIRS rating, then the discernability of certain features is possible. For example, in many of the MSTAR images, the ground sampling distance (GSD) resolution is 1 ft (0.3 m), which corresponds to a NIIRS of 7.

NIIRS	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
Resolution	> 9m	4.5-9.0m	2.5-4.5m	1.2-2.5m	0.75-1.2m	0.4-0.75m	0.2-0.4m	0.1-0.2 m	<0.1m
Detect	Road networks	Defense area	Buildings	Convoy	Semi-Truck	Wheeled vs. tracked tanks	Medium tank versus car	Turrets	Guns

The SAR NIIRS and resolution (GSD)[**47**] is then determined by sensor analogy to a general image quality equation (GIQE-4):

$$\boxed{NIIRS_{IR} = 10.751 - A * \log_{10}(GSD) + B * \log_{10}(RER) - 0.656\,H - [0.344(G/SNR)]}$$

where *GSD* is the geometric mean of the ground sample distance, *H* is the geometric mean height due to edge overshoot, *RER* is the geometric mean of the normalized relative edge response, *G* is the noise gain, *SNR* is the signal to noise ratio, *A* is constant (3.32 if *RER* = 0.9, 3.16 if *RER*<0.9), and *B* is constant (1.559 if *RER* = 0.9, 2.817 if *RER*< 0.9).

Typically, the SAR GSD is related to the $NIIRS_{IR}$ via:

$$GSD = 10^{[(10.751 - NIIRSIR)/A]}$$

R. Driggers, J. Ratches, J. Leachtenauer, R. Kistner, “Synthetic aperture radar target acquisition model based on a National Imagery Interpretability Rating Scale to probability of discrimination conversion”, *Optical Engineering*, 42(7), July 2003.

© Majumder, Blasch, Garren

National Interpretability Imagery Rating Scale (NIIRS): Example

national imagery interpretability rating scale (NIIRS) to assess the quality of SAR data. Given a SAR image with a NIIRS rating, then the discernability of certain features is possible. For example, in many of the MSTAR images, the ground sampling distance (GSD) resolution is 1 ft (0.3 m), which corresponds to a NIIRS of 7.

NIIRS	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
Resolution	> 9m	4.5-9.0m	2.5-4.5m	1.2-2.5m	0.75-1.2m	0.4-0.75m	0.2-0.4m	0.1-0.2 m	<0.1m
Detect	Road networks	Defense area	Buildings	Convoy	Semi-Truck	Wheeled vs. tracked tanks	Medium tank versus car	Turrets	Guns

Through experiment, an **improved empirically fit result** was determined to be:

$$\boxed{NIIRS_{IR} = 1.14 + 0.18*NIIRS_{SAR} + 0.08*NIIRS^2_{SAR}}$$

Using $y = ax^2 + bx + c$, then the solution is:

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{(2a)}$$

$$NIIRS_{SAR} = \frac{-0.18 \pm \sqrt{(0.18)^2 - 4\,(0.08)\,(1.14 - NIIRS_{IR})}}{(2 * 0.08)}$$

Knowing that NIIRS > 1, then the following are determined as:

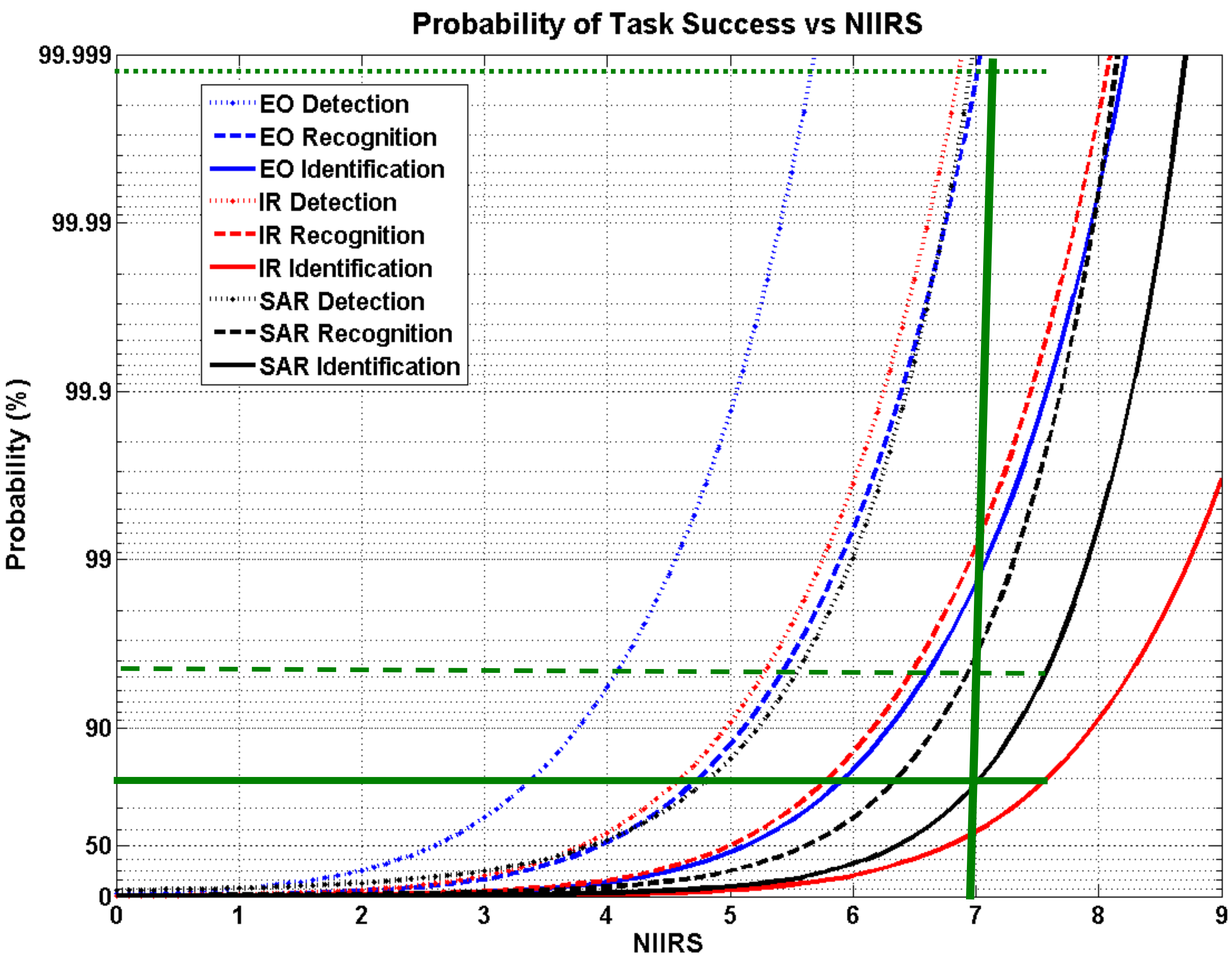
NIIRS-I	NIIRS - S	GSD (m)	2*GSD(m)
2	2.3	5.88	11.76
3	3.8	2.84	5.67
4	5.0	1.37	2.74
5	5.9	0.66	1.32
6	6.8	0.32	0.64
7	**7.5**	**0.15**	**0.31**
8	8.2	0.07	0.15
9	8.9	0.04	0.07

E. Blasch, H-M. Chen, J. M. Irvine, Z. Wang, G. Chen, J. Nagy, S. A. Scott, "Prediction of compression induced image interpretability degradation," *Opt. Eng.*57(4), 043108, 2018.

© Majumder, Blasch, Garren

SAR NIIRS Analysis

Figure 9.19



SAR

Detection 99%

Recognition 92%

Identification/Classification 80%

B. Kahler and E. Blasch, "Predicted Radar/Optical Feature Fusion Gains for Target Identification," *Proc. IEEE Nat. Aerospace Electronics Conf (NAECON)*, 2010.

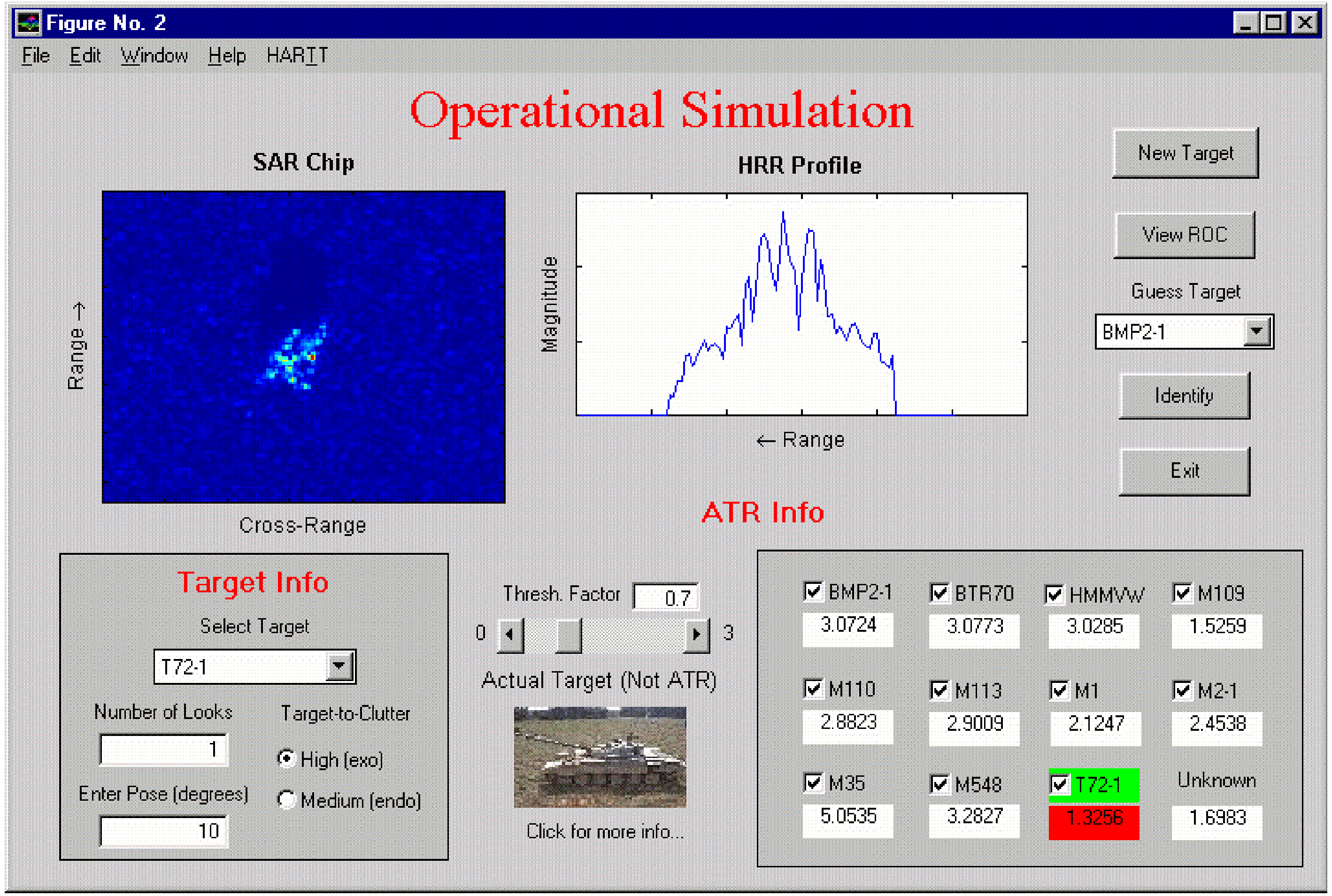
© Majumder, Blasch, Garren

User Interface

E. P. Blasch "Assembling a distributed fused Information-based Human-Computer Cognitive Decision Making Tool," *IEEE Aerospace and Electronic Systems Magazine,* Vol. 15, No. 5, pp. 11-17, May 2000.

- **Constructive Test**

  simulation involving
  simulated people operating
  simulated systems

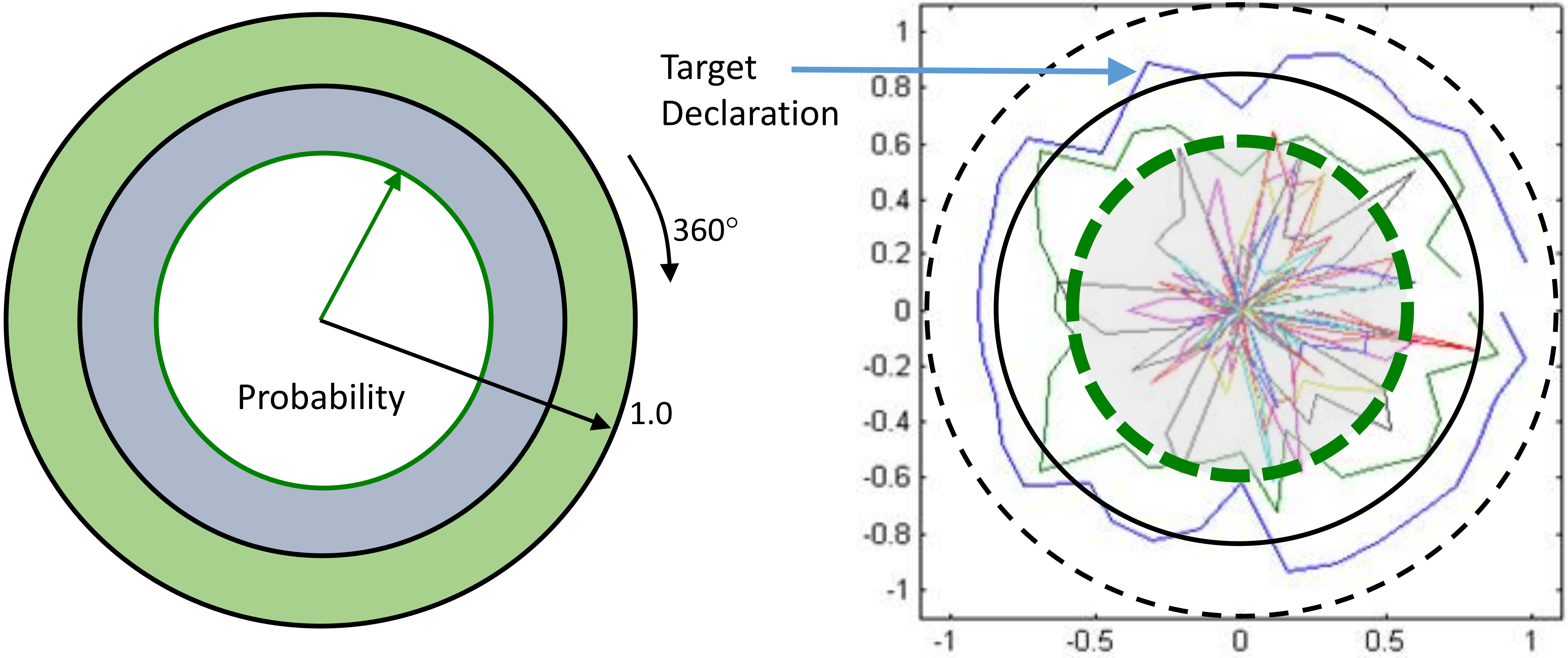


E. Blasch, *Derivation of a Belief Filter for Simultaneous High Range Resolution Radar Tracking and Identification,* Ph.D. Thesis, Wright State University, 1999.

Display of Results to User

- **Visualization of metrics (360° around target)**

Figure 9.20



E. P. Blasch and P. Svenmarck, "Target recognition using Vehicle Separation plots (VSP) for human assessment," *5th World Multi-conference On Systems, Cybernetics, and Informatics (SCI 2001)*, July 2001.

© Majumder, Blasch, Garren

User Interface (for deployment)

Figure 9.21

- **Live Test**

  simulation involving

  real people operating

  simulated systems

- **Virtual Test**

  simulation involving

  real people operating

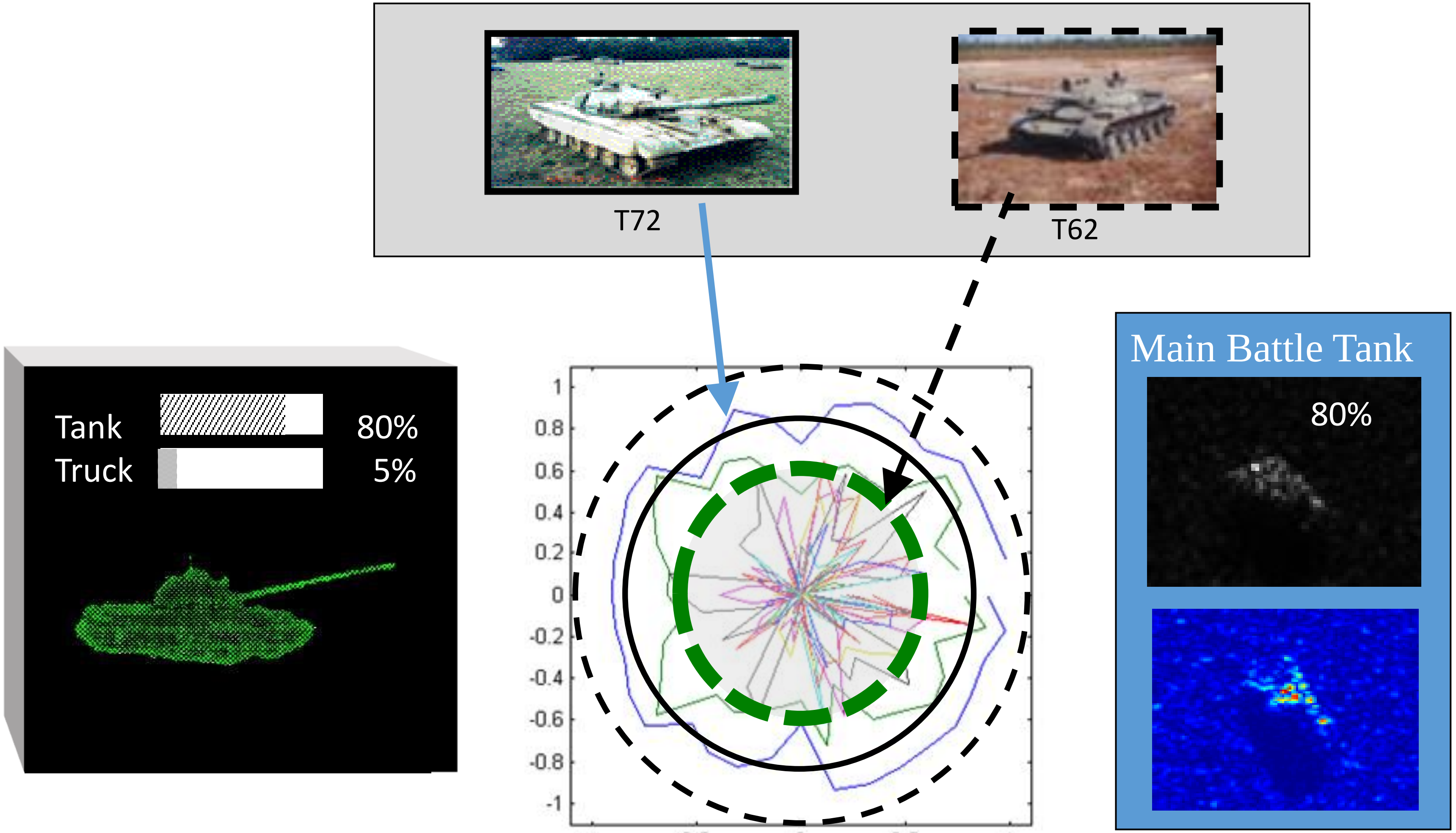
  simulated systems

- **Constructive Test**

  simulation involving

  simulated people operating

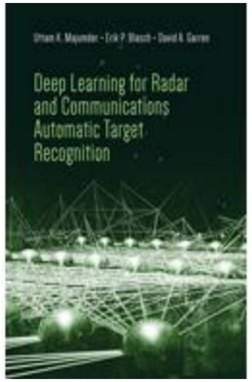
  simulated systems



ATR Score **ATR Comparison** **ATR Feature**

© Majumder, Blasch, Garren

Chapter 9 Outline



© Majumder, Blasch, Garren