AUSTRALIAN CENTRE FOR ROBOTIC VISION

Neural Algebra of Classifiers

Rodrigo Santa Cruz, Basura Fernando, Anoop Cherian, and Stephen Gould

Australian National University

1 - INTRODUCTION & MOTIVATION

- Which one is the albatross?

Albatrosses are birds with hooked beak and large wingspan. (hooked beak AND large wingspan)

- Which ones is the frigatebird?

Frigatebirds seem black albatrosses with white or red pouch. (albatross AND (white pouch OR red pouch))

"The **human recognition system** is fundamentally **compositional**, so the **unseen complex concepts** are recognized from the composition of **simple visual primitives** according to **well-defined rules**."

2 - GOALS & CONTRIBUTIONS

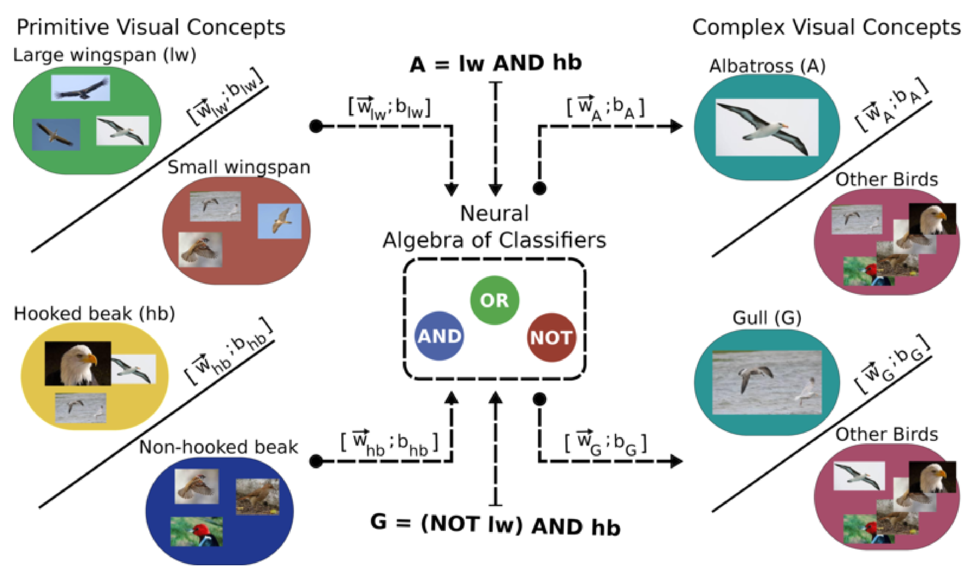
We aim to develop a learning framework to synthesize classifiers for arbitrary compositions of visual primitives.

Contributions:

- Learning framework for composition of classifiers.
- Compositional neural network based model which minimizes the classification error of a subset of visual compositions and generalizes for unseen compositions.
- The proposed framework can recognize unseen classes, subclasses and specific instances of objects without additional annotation effort.

3 - PROBLEM FORMULATION

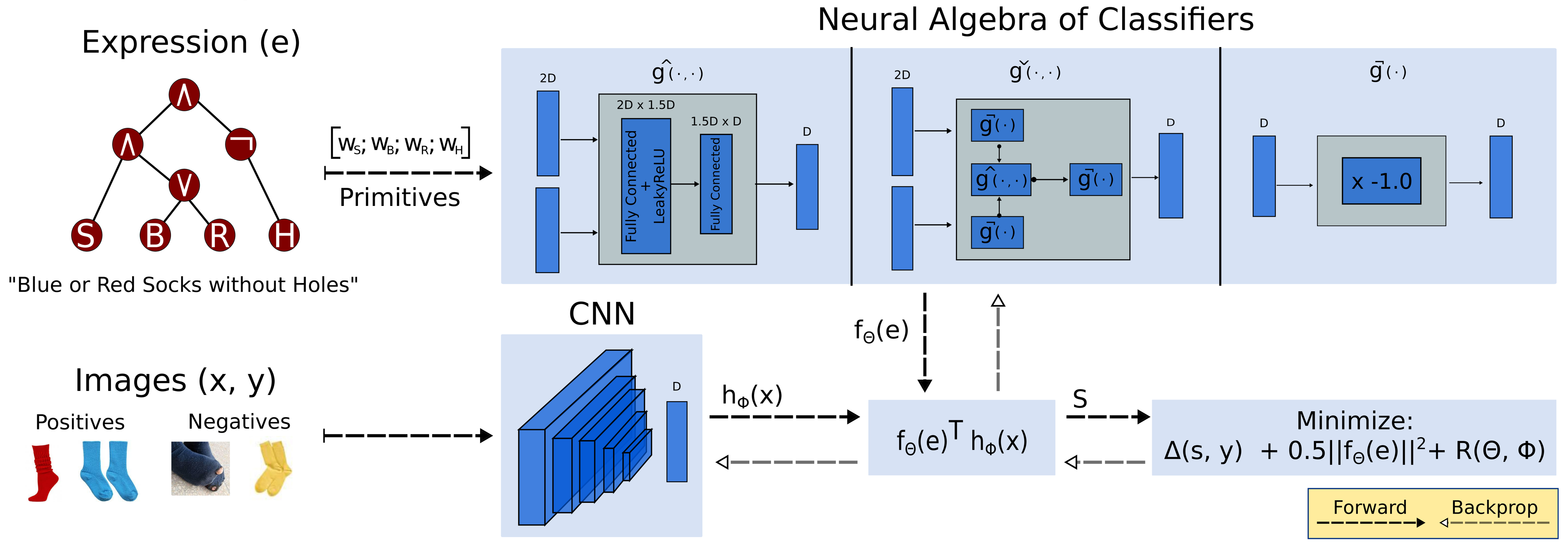
We propose an **algebra of visual primitives**:



- Visual primitives (p): known simple visual concepts, e.g., hooked beak (hb) and large wingspan (lw).
- Composition rules: (∧, AND) conjunction, (∨, OR) disjunction, and (¬, NOT) negation.
- Expressions (e): visual concepts expressed as multiple compositions of primitives and composition rules, e.g., albatrosses = hb AND lw.

4 - LEARNING & INFERENCE

We propose to learn a parameterized function ($w_e = f_\Theta(e)$) that maps the space of expressions to the space of binary classifiers using a relative small subset of training expressions and relying on the classifier similarity to generalize for unknown expressions.



Baselines:

- **Supervised:** SVMs trained to discriminate images according to training expressions.
- **Independent Classifiers:**
  - P(a AND b) = P(a) x P(b)
  - P(a OR b) = P(a) + P(b) - P(a AND b)
  - P(NOT a) = 1 - P(a)

5 - COMPOSITION FUNCTIONS

We model the function $f_\Theta(e)$ as a set of composition functions by decomposing expressions in simple terms and evaluating them recursively.

$$g_\theta^{\wedge}(w_a, w_b) = \text{Neural Network}(w_a, w_b)$$

$$g^{\neg}(w) = -w$$

$$g^{\vee}(w_a, w_b) = g^{\neg}(g^{\wedge}(g^{\neg}(w_a), g^{\neg}(w_b)))$$

- These composition functions are autoregressive models.
- The negation is defined analytically.
- The disjunction is defined according to the **De Morgan's Laws**.
- The conjunction is a **multilayer perceptron (MLP)**.

6 - RESULTS

Simple Binary Expressions:

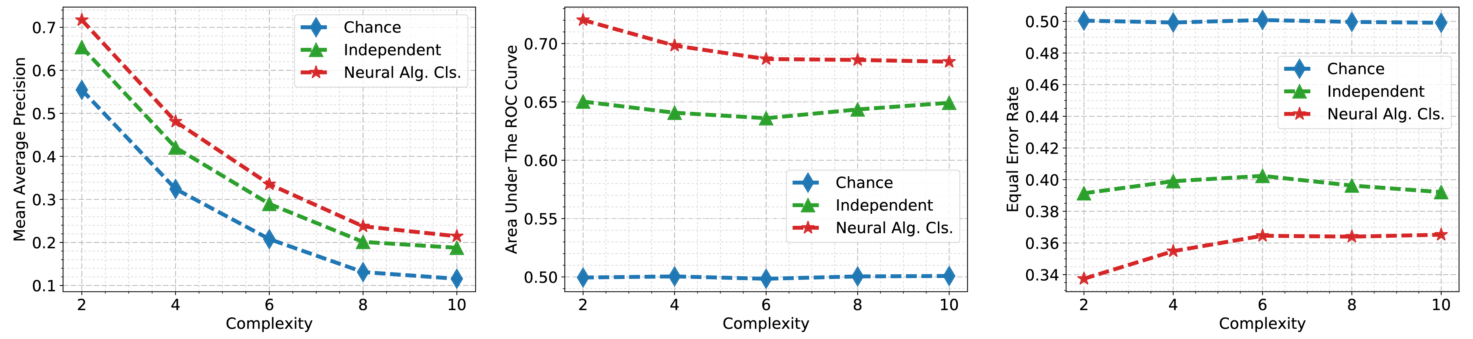
Table 1. Evaluating known/unknown disjunctive and conjunctive expressions on the CUB-200 Birds dataset.

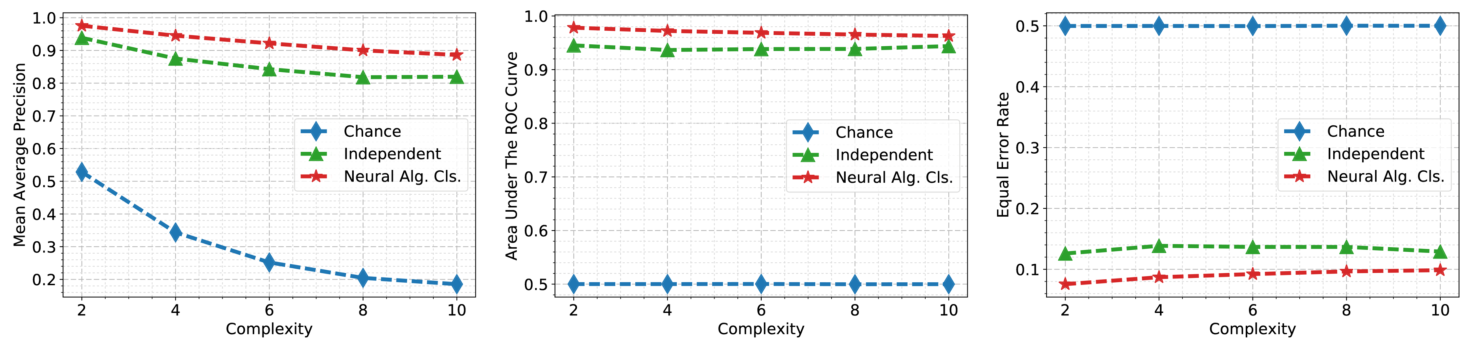
	Disjunctive Expressions						Conjunctive Expressions					
	Known Exp.			Unknown Exp.			Known Exp.			Unknown Exp.		
Metrics	MAP	AUC	EER	MAP	AUC	EER	MAP	AUC	EER	MAP	AUC	EER
Chance	39.70	50.00	50.0	40.60	50.00	50.0	4.55	50.0	50.0	4.59	50.0	50.0
Supervised	65.25	74.76	31.58	-	-	-	22.87	78.02	29.69	-	-	-
Independent	58.73	68.39	36.76	60.66	69.28	36.10	17.23	77.22	29.94	19.16	78.00	29.28
Neural Alg. Classifiers	**70.10**	**77.36**	**29.44**	**71.18**	**77.76**	**29.04**	**23.09**	**81.54**	**26.36**	**23.87**	**81.98**	**25.85**

Table 2. Evaluating known/unknown disjunctive and conjunctive expressions on the AwA2 dataset.

	Disjunctive Expressions						Conjunctive Expressions					
	Known Exp.			Unknown Exp.			Known Exp.			Unknown Exp.		
Metrics	MAP	AUC	EER	MAP	AUC	EER	MAP	AUC	EER	MAP	AUC	EER
Chance	53.19	50.0	50.0	53.04	50.0	50.0	18.77	50.0	50.0	21.17	50.0	50.0
Supervised	97.47	97.20	8.13	-	-	-	94.90	98.53	6.00	-	-	-
Independent	97.28	97.12	8.70	97.86	97.58	6.77	93.95	98.13	6.80	93.90	97.87	7.36
Neural Alg. Classifiers	**98.84**	**98.67**	**5.84**	**99.05**	**98.91**	**5.24**	**95.95**	**98.79**	**5.29**	**96.50**	**98.81**	**5.34**

Complex Unknown Expressions: $(p_1 \vee q_1) \wedge (p_2 \vee q_2) \wedge \dots \wedge (p_c \vee q_c)$





Qualitative Evaluation