



دانشگاه صنعتی شریف  
دانشکده مهندسی کامپیوتر

پایان نامه کارشناسی ارشد  
هوش مصنوعی

دسته بندی ریزدانه ای تصاویر

نگارش

یاسر سوری

استاد راهنما

دکتر شهره کسایی

بهار ۱۳۹۴

تقديم به

به نام او

دانشگاه صنعتی شریف  
دانشکده مهندسی کامپیوتر

پایان نامه کارشناسی ارشد

عنوان: دسته بندی ریزدانه ای تصاویر

نگارش: یاسر سوری

کمیته داوران

استاد راهنما: دکتر شهره کسایی  
امضاء: . . . . .

ممتحن داخلی: دکتر محمد تقی منظوری شلمانی  
امضاء: . . . . .

داور خارجی: دکتر بابک نجار اعرابی  
امضاء: . . . . .

تاریخ: . . . . .

## قدردانی

باعث افتخار بود که

## چکیده

دسته‌بندی تصویر ریزدانه‌ای عبارت است از

**کلمات کلیدی:** بینایی کامپیوتری، بازشناسی شیء، یادگیری عمیق، دسته‌بندی تصاویر، بازشناسی ریزدانه‌ای.

# فهرست مطالب

دوم	فهرست شکل‌ها
سوم	فهرست جدول‌ها
۱	فصل ۱: مقدمه
۳	فصل ۲: روش‌های پیشین
۳	۱-۲ روش‌های دسته‌بندی تصویر
۴	فصل ۳: روش پیشنهادی
۵	فصل ۴: نتایج تجربی
۵	۱-۴ معیار ارزیابی
۵	۲-۴ نتایج نهایی روش پیشنهادی
۷	فصل ۵: جمع‌بندی و کارهای آتی
۷	۱-۵ جمع‌بندی
۷	۲-۵ کارهای آینده
۷	۱-۲-۵ روشی برای
۸	مراجع

## فهرست شکل‌ها

شکل ۱-۱	نمونه‌هایی از انواع تنوع در تصاویر طبیعی که برای بازشناسی باید مورد توجه
۲	قرار بگیرد . . . . .

## فهرست جدول‌ها

۱-۴ دقت روش‌های نهایی پیشنهادی (مشخص شده توسط \*) در مقایسه با روش مرز دانش. برای روش‌های پیشنهادی میانگین و انحراف معیار در سه آزمایش گزارش شده است. ۶ . . . . .



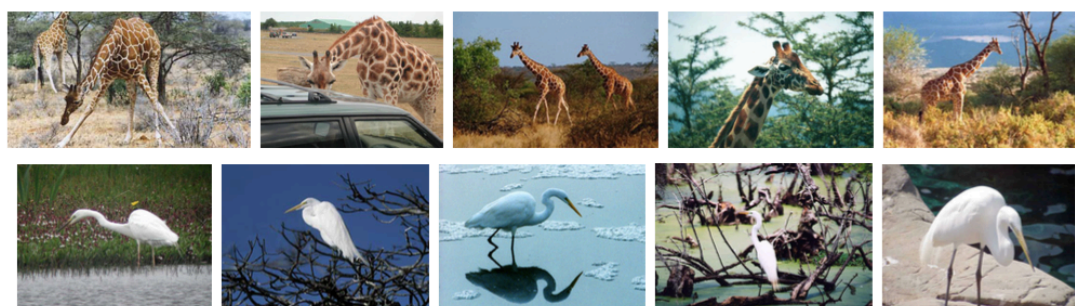
---

# فصل ۱

---

## مقدمه

البته چالش اصلی این مسئله تنوع تصاویر طبیعی است. این تنوع حاصل تفاوت در حالت <sup>۱</sup> است. می توانید به شکل ۱-۱ ارجاع دهید.



(آ)



(ب)

شکل ۱-۱: نمونه‌هایی از انواع تنوع در تصاویر طبیعی که برای بازشناسی باید مورد توجه قرار بگیرد [۱].

(آ) اشیاء معمولاً به خاطر پوشیدگی و شلوغی به سختی قابل بازشناسی هستند. حیوانات از جمله اشیاء غیر صلب هستند و در حالات مختلفی ظاهر می‌شوند. (ب) تصاویر صحنه‌های طبیعی مانند کوهستان دارای تفاوت درون دسته‌ای بسیار زیادی هستند. همچنین برخی تصاویر از محیط‌های داخلی (مانند آشپزخانه و اتاق نشیمن) به سختی از هم قابل تمایز هستند.

---

## فصل ۲

---

### روش‌های پیشین

در سال‌های اخیر پیشرفت‌های بسیار چشم‌گیری در زمینه‌ی دسته‌بندی تصویر رخ داده است.

#### ۱-۲ روش‌های دسته‌بندی تصویر

همانطور که بیان شد، در دسته‌بندی تصویر، باید تصویر را با توجه به محتوایش دسته‌بندی کنیم.

---

## فصل ۳

---

### روش پیشنهادی

روش پیشنهادی ...

---

## فصل ۴

---

### نتایج تجربی

در این بخش به نتایج تجربی حاصل از پیاده‌سازی روش‌های معرفی شده در فصل ۳ خواهیم پرداخت.

#### ۱-۴ معیار ارزیابی

فرض کنید که داده‌های آزمایشی ما به صورت زوج مرتب‌های  $\{(x_i, y_i)\}_{i=1}^{N_{test}}$  موجود باشد. که در آن  $x_i$  تصویر،  $y_i$  برچسب مربوط به آن و  $N_{test}$  تعداد تصاویر آزمایشی است. همچنین در پایگاه داده پرندگان کلتک  $N_{test}$  برابر با ۵۷۹۴ است.

حال برای ارزیابی تابع دسته‌بند یعنی  $f(x)$  از فرمول

$$mA = \frac{1}{C} \sum_{c=1}^C \frac{1}{|\alpha(c)|} \sum_{i \in \alpha(c)} \mathbb{I}(f(x_i) = y_i) \quad (1-4)$$

دقت میانگین را محاسبه می‌کنیم.

#### ۲-۴ نتایج نهایی روش پیشنهادی

حالت دوم موسوم به DeepRF(All) که در آن روش پیشنهاد شده به غیر از مستطیل محیطی سر و بدن پرنده، لازم است که مستطیل محیطی کل پرنده را نیز تخمین بزند. به دلیل استفاده از جنگل تصادفی در این روش و ذات تصادفی بودن آن، آزمایش‌ها را سه بار انجام داده‌ایم و میانگین و انحراف

جدول ۴-۱: دقت روش‌های نهایی پیشنهادی (مشخص شده توسط \*) در مقایسه با روش مرز دانش. برای روش‌های پیشنهادی میانگین و انحراف معیار در سه آزمایش گزارش شده است.

سال	نام	آموزش		آزمایش		دقت میانگین دسته‌بندی (درصد)
		پنجره محیطی	مکان اجزا	پنجره محیطی	مکان اجزا	
۲۰۱۴	PRCNN [۲]	✓	✓	✓		۷۶/۳۷
-	* DeepRF	✓	✓	✓		۷۳/۷۸ (۰/۳۲±)
۲۰۱۴	PRCNN [۲]	✓	✓			۷۳/۸۹
-	* DeepRF(All)	✓	✓			۷۲/۰۲ (۰/۳۳±)

معیار دقت را گزارش خواهیم کرد.

همانطور که در جدول ۴-۱ نیز ذکر شده است، روش پیشنهادی موسوم به DeepRF به دقت میانگین ۷۳/۷۸ دست پیدا می‌کند که قابل مقایسه با روش مرز دانش با دقت میانگین ۷۶/۳۷ است. همچنین روش پیشنهادی موسوم به DeepRF(All) به دقت میانگین ۷۲/۰۲ دست پیدا می‌کند. روش مرز دانش نیز به دقت میانگین ۷۳/۸۹ درصد در حالتی که مستطیل محیطی در زمان آزمایش در دسترس نیست دست پیدا می‌کند.

به طور کلی می‌توان گفت که روش پیشنهاد شده از نظر کارایی بسیار سریع‌تر از روش مرز دانش است، از نظر سادگی بسیار ساده‌تر از آن است و از نظر کاربردی در موارد مختلف دیگری نیز با کمترین تغییرات می‌تواند مورد استفاده قرار بگیرد. با این حال از نظر دقت میانگین دسته‌بندی نتیجه‌ای بسیار نزدیک نسبت به آن کسب می‌کند.

---

## فصل ۵

---

### جمع‌بندی و کارهای آتی

#### ۱-۵ جمع‌بندی

در این پایان‌نامه به بررسی مسئله دسته‌بندی ریزدانه‌ای و اهمیت آن و همچنین روش‌های موجود برای حل آن پرداختیم. سپس دو روش جدید برای حل این مسئله بر روی پایگاه داده پرندگان کلتک پیشنهاد دادیم.

#### ۲-۵ کارهای آینده

##### ۱-۲-۵ روشی برای

یکی از محدودیت‌های روش پیشنهادی ...

## مراجع

- [1] S. Lazebnik, “Local, semi-local and global models for texture, object and scene recognition,” *Ph.D. Dissertation, University of Illinois at Urbana-Champaign*, 2006.
- [2] N. Zhang, J. Donahue, R. Girshick, and T. Darrell, “Part-based r-cnns for fine-grained category detection,” in *ECCV*. Springer, 2014.



## **Abstract**

Fine-grained image classification is

**Keywords:** Computer vision, Object recognition, Deep learning, Image classification, Fine-grained recognition.



Sharif University of Technology  
Department of Computer Engineering

**M.Sc. Thesis**  
**Artificial Intelligence**

**Fine-grained Image Classification**

By

**Yaser Souri**

Supervisor

**Prof. Shohreh Kasaei**

Spring 2015