

یادگیری عمیق در بینای کامپیوتر

مؤلف

محمد القندی

مترجم

ایوب ترکیان

نیاز دانش

فهرست مطالب

شماره صفحه

عنوان

۱۱	فصل ۱ / مقدمه بینایی کامپیوتر
۱۲	۱.۱ شهود بینایی کامپیوتر
۱۳	۱.۱.۱ ادراک عینی
۱۳	۲.۱.۱ سیستم‌های بینایی
۱۳	۳.۱.۱ سیستم‌های بینایی انسان
۱۵	۴.۱.۱ سیستم‌های بینایی AI
۱۵	۵.۱.۱ وسیله حس کردن
۱۷	۶.۱.۱ وسیله تفسیر کردن
۱۸	۷.۱.۱ یادگیری ماشین و مغز انسان
۱۹	۲.۱ کاربردهای بینایی کامپیوتر
۲۰	۱.۲.۱ دسته‌بندی تصویر
۲۲	۲.۲.۱ آشکارسازی شیء و جانمایی
۲۲	۳.۲.۱ توضیح تصویر اتوماتیک
۲۳	۴.۲.۱ تولید تصویر (انتقال سبک)
۲۴	۵.۲.۱ خلق تصویر
۲۵	۶.۲.۱ نتیجه‌گیری
۲۶	۳.۱ خط لوله بینایی کامپیوتر-تصویر کلان
۲۹	۴.۱ تصویر ورودی
۲۹	۱.۴.۱ تصویر به عنوان تابع
۳۱	۲.۴.۱ مشاهده تصویر توسط کامپیوتر
۳۱	۳.۴.۱ تصاویر رنگی
۳۴	۵.۱ پیش‌پردازش تصویر
۳۴	۱.۵.۱ پردازش تصویر
۳۴	۲.۵.۱ چرا پیش‌پردازش تصویر
۳۸	۶.۱ استخراج ویژگی
۳۸	۱.۶.۱ ویژگی در بینایی کامپیوتر
۴۰	۲.۶.۱ ویژگی مناسب (مفید)
۴۳	۳.۶.۱ استخراج ویژگی‌ها (دستی در مقابل اتوماتیک)
۴۳	۴.۶.۱ ویژگی‌های دستی در یادگیری ماشین سنتی
۴۳	۵.۶.۱ استخراج اتوماتیک ویژگی در یادگیری عمیق
۴۴	۶.۶.۱ استفاده از ویژگی
۴۶	۷.۱ الگوریتم یادگیری طبقه‌گر
۴۷	۸.۱ خلاصه فصل

فصل ۲ / یادگیری عمیق و شبکه‌های عصبی

۵۲	۱.۲ ایده پرسپترون
۵۳	۱.۱.۲ پرسپترون چیست؟

۵۷	یادگیری پرسپترون	۲.۱.۲
۵۸	کفایت یک نورون	۳.۱.۲
۶۰	پرسپترون چندلایه (MLP)	۲.۲
۶۱	معماری پرسپترون چندلایه	۱.۲.۲
۶۲	لایه‌های پنهان	۲.۲.۲
۶۳	تعداد لایه و گره	۳.۲.۲
۶۵	موارد مهم MLP	۴.۲.۲
۶۵	توابع فعال‌سازی	۳.۲
۶۷	تابع انتقال خطی	۱.۳.۲
۶۸	طبقه‌گر دوتایی	۲.۳.۲
۶۹	تابع سیگمایی/الجیستیک	۳.۳.۲
۷۱	تابع سافت‌مکس	۴.۳.۲
۷۲	تابع تانژانت سهموی (tanh)	۵.۳.۲
۷۳	واحد خطی همسوسده	۶.۳.۲
۷۴	ReLU نشستی	۷.۳.۲
۷۷	پیش‌خور	۴.۲
۷۸	محاسبات پیش‌خور	۱.۴.۲
۷۹	یادگیری ویژگی	۲.۴.۲
۸۳	توابع خطا	۵.۲
۸۳	تابع خطا	۱.۵.۲
۸۴	ضرورت تابع خطا	۲.۵.۲
۸۴	مثبت بودن خطا	۳.۵.۲
۸۵	خطای مربع میانگین	۴.۵.۲
۸۶	انتروپی متقاطع	۵.۵.۲
۸۸	یادداشت نهایی	۶.۵.۲
۸۹	الگوریتم‌های بهینه‌سازی	۶.۲
۹۰	بهینه‌سازی چیست؟	۱.۶.۲
۹۳	فرود گرادیان ناپیوسته	۲.۶.۲
۱۰۰	فرود گرادیان استوکاستیک	۳.۶.۲
۱۰۱	فرود گرادیان مینی‌بچ	۴.۶.۲
۱۰۱	نکات فرود گرادیان	۵.۶.۲
۱۰۳	پس‌انتشار	۷.۲
۱۰۴	پس‌انتشار چیست؟	۱.۷.۲
۱۰۷	نکات پس‌انتشار	۲.۷.۲
۱۰۷	خلاصه و نکات	۸.۲

فصل ۳ / شبکه‌های عصبی کانولوشن

۱۱۲	طبقه‌بندی تصویر با MLP	۱.۳
۱۱۳	لایه ورودی	۱.۱.۳
۱۱۶	لایه‌های پنهان	۲.۱.۳
۱۱۶	لایه خروجی	۳.۱.۳
۱۱۷	تلفیق	۴.۱.۳
۱۱۹	نواقص MLP-ها در پردازش تصویر	۵.۱.۳
۱۲۲	نتیجه‌گیری	۶.۱.۳

۱۲۳	معماری CNN-ها	۲.۳
۱۲۳	تصویر کلان	۱.۲.۳
۱۲۶	بررسی دقیق تر استخراج ویژگی	۲.۲.۳
۱۲۷	بررسی دقیق تر طبقه بندی	۳.۲.۳
۱۲۸	مؤلفه های بنیادی CNN	۳.۳
۱۲۹	لایه های کانولوشن (CONV)	۱.۳.۳
۱۳۸	لایه های رأی گیری یا فرونمونه برداری (POOL)	۲.۳.۳
۱۴۱	علت استفاده از لایه رأی گیری	۳.۳.۳
۱۴۴	لایه های تمام وصل (FC)	۴.۳.۳
۱۴۶	طبقه بندی تصویر با CNN-ها	۴.۳
۱۴۷	ساخت معماری مدل	۱.۴.۳
۱۴۷	خلاصه مدل	۲.۴.۳
۱۴۹	تعداد پارامترها (اوزان)	۳.۴.۳
۱۵۱	افزودن لایه دورریز	۵.۳
۱۵۱	فرابرازش	۱.۵.۳
۱۵۲	لایه دورریز	۲.۵.۳
۱۵۳	ضرورت لایه های دورریز	۳.۵.۳
۱۵۵	محل لایه دورریز	۴.۵.۳
۱۵۶	کانولوشن تصاویر رنگی	۶.۳
۱۶۲	خلاصه فصل	۷.۳
۱۶۴	پروژه: دسته بندی تصاویر برای تصاویر رنگی	۸.۳
۱۶۴	بارگذاری مجموعه داده ها	۱.۸.۳
۱۶۴	پیش پردازش تصویر	۲.۸.۳
۱۶۸	تعریف معماری مدل	۳.۸.۳
۱۷۱	کامپایل کردن مدل	۴.۸.۳
۱۷۲	آموزش مدل	۵.۸.۳
۱۷۴	بارگذاری مدل با بهترین val_acc	۶.۸.۳
۱۷۴	ارزیابی مدل	۷.۸.۳

فصل ۴ / ساختاردهی پروژه ها و تنظیم فرآیندها

۱۷۸	تعریف متریک های عملکرد	۱.۴
۱۷۹	صحت	۱.۱.۴
۱۷۹	ماتریس درهم ریختگی	۲.۱.۴
۱۸۰	دقت و به خاطر آوری	۳.۱.۴
۱۸۱	نمره F	۴.۱.۴
۱۸۲	تعریف مدل خط مینا	۲.۴
۱۸۴	آماده سازی داده ها	۳.۴
۱۸۴	تقسیم داده ها	۱.۳.۴
۱۸۷	پیش پردازش داده ها	۲.۳.۴
۱۹۰	ارزیابی مدل	۴.۴
۱۹۰	تشخیص فرابرازش و فروبرازش	۱.۴.۴
۱۹۲	رسم نمودارهای یادگیری	۲.۴.۴
۱۹۴	تمرین شبکه کوچک	۳.۴.۴
۱۹۷	بهبود شبکه	۵.۴

۱۹۷	جمع‌آوری داده یا تنظیم	۱.۵.۴
۱۹۸	پارامتر و فراپارامتر	۲.۵.۴
۱۹۹	فراپارامترهای شبکه عصبی	۳.۵.۴
۲۰۰	معماری شبکه عصبی	۴.۵.۴
۲۰۲	یادگیری و بهینه‌سازی	۵.۵.۴
۲۱۶	تکنیک‌های تنظیم	۶.۵.۴
۲۲۰	۶.۴ نرمالسازی بیچ (BN)	
۲۲۰	مسئله شیفت کوواریانس	۱.۶.۴
۲۲۲	شیفت کوواریانس در شبکه‌های عصبی	۲.۶.۴
۲۲۳	نحوه کار نرمال‌سازی بیچ	۳.۶.۴
۲۲۴	نرمال‌سازی بیچ در کراس	۴.۶.۴
۲۲۵	جمع‌بندی نرمال‌سازی بیچ	۵.۶.۴
۲۲۵	۷.۴ خلاصه فصل	
۲۲۵	سخت‌ن پایداری	۱.۷.۴
۲۲۶	نکات تنظیم فراپارامترها	۲.۷.۴
۲۲۷	نکات مبانی یادگیری عمیق	۳.۷.۴
۲۲۸	گام بعدی	۴.۷.۴
۲۲۸	توصیه مؤلف	۵.۷.۴
۲۲۸	۸.۴ پروژه متریک‌های عملکرد	
۲۲۹	وارد کردن وابستگی‌ها	۱.۸.۴
۲۲۹	آماده‌سازی داده‌ها	۲.۸.۴
۲۳۱	ساخت معماری مدل	۳.۸.۴
۲۳۳	آموزش مدل	۴.۸.۴
۲۳۴	ارزیابی مدل	۵.۸.۴
۲۳۴	رسم نمودارهای یادگیری	۶.۸.۴
۲۳۵	بهبودهای اضافی	۷.۸.۴

فصل ۵ / معماری‌های CNN پیشرفته ۲۳۷

۲۴۰	۱.۵ الگوهای طراحی CNN	
۲۴۰	۱.۱.۵ الگوی شماره ۱	
۲۴۰	۲.۱.۵ الگوی شماره ۲	
۲۴۱	۳.۱.۵ الگوی شماره ۳	
۲۴۲	۲.۵ LeNet-5	
۲۴۲	۱.۲.۵ معماری LeNet-5	
۲۴۳	۲.۲.۵ LeNet-5 در کراس	
۲۴۴	۳.۲.۵ برپایش فراپارامترهای یادگیری	
۲۴۵	۴.۲.۵ عملکرد LeNet-5 روی MNIST	
۲۴۶	۳.۵ AlexNet	
۲۴۸	۱.۳.۵ معماری AlexNet	
۲۴۸	۲.۳.۵ ویژگی‌های ابداعی AlexNet	
۲۴۹	۱.۲.۳.۵ تابع فعال‌سازی ReLU	
۲۴۹	۲.۲.۳.۵ لایه دورریزی	
۲۵۰	۳.۲.۳.۵ تقویت داده‌ها	
۲۵۰	۴.۲.۳.۵ نرمال‌سازی پاسخ محلی	

۲۵۰	تنظیم وزن	۵.۲.۳.۵
۲۵۰	آموزش روی چند GPU	۶.۲.۳.۵
۲۵۱	AlexNet در کراس	۳.۳.۵
۲۵۴	برپایش فرآپارامترهای یادگیری	۴.۳.۵
۲۵۵	عملکرد AlexNet روی CIFAR	۵.۳.۵
۲۵۶	VGGNet	۴.۵
۲۵۶	ویژگی‌های ابداعی VGGNet	۱.۴.۵
۲۵۸	پیکربندی VGGNet	۲.۴.۵
۲۵۹	VGG-16 در کراس	۳.۴.۵
۲۶۰	فرآپارامترهای یادگیری	۴.۴.۵
۲۶۱	عملکرد VGGNet روی CIFAR	۵.۴.۵
۲۶۱	Inception و GoogLeNet	۵.۵
۲۶۱	ویژگی‌های ابداعی Inception	۱.۵.۵
۲۶۳	مدول Inception - نسخه ساده‌انگارانه	۲.۵.۵
۲۶۴	مدول Inception با کاهش بُعدیت	۳.۵.۵
۲۶۸	معماری Inception	۴.۵.۵
۲۷۱	GoogLeNet در کراس	۵.۵.۵
۲۷۵	فرآپارامترهای یادگیری	۶.۵.۵
۲۷۶	عملکرد Inception روی CIFAR	۷.۵.۵
۲۷۶	ResNet	۶.۵
۲۷۶	ویژگی‌های ابداعی ResNet	۱.۶.۵
۲۷۹	بلوک‌های باقیمانده	۲.۶.۵
۲۸۲	پیاده‌سازی ResNet در کراس	۳.۶.۵
۲۸۵	فرآپارامترهای یادگیری	۴.۶.۵
۲۸۵	عملکرد ResNet روی CIFAR	۵.۶.۵
۲۸۶	خلاصه فصل	۷.۵

فصل ۶ / یادگیری انتقالی ۲۸۹

۲۹۰	مسائل حل شده با یادگیری انتقالی	۱.۶
۲۹۲	یادگیری انتقالی	۲.۶
۳۰۰	نحوه کار یادگیری انتقالی	۳.۶
۳۰۲	نحوه یادگیری ویژگی‌ها توسط شبکه عصبی	۱.۳.۶
۳۰۴	قابلیت انتقال ویژگی‌ها	۲.۳.۶
۳۰۵	رویکردهای یادگیری انتقالی	۴.۶
۳۰۵	شبکه از پیش‌آموزش دیده به‌عنوان طبقه‌گر	۱.۴.۶
۳۰۷	شبکه از پیش‌آموزش دیده به‌عنوان استخراجگر ویژگی	۲.۴.۶
۳۰۹	تدقیق	۳.۴.۶
۳۱۲	انتخاب سطح مناسب یادگیری انتقالی	۵.۶
۳۱۳	سناریوی ۱: مجموعه داده هدف کوچک و مشابه	۱.۵.۶
۳۱۳	سناریوی ۲: مجموعه داده هدف بزرگ و مشابه	۲.۵.۶
۳۱۴	سناریوی ۳: مجموعه داده هدف کوچک و متفاوت	۳.۵.۶
۳۱۴	سناریوی ۴: مجموعه داده هدف بزرگ و متفاوت	۴.۵.۶
۳۱۴	جمع‌بندی سناریوهای یادگیری انتقالی	۵.۵.۶
۳۱۵	مجموعه داده‌های منبع باز	۶.۶

۳۱۶	MNIST	۱.۶۶
۳۱۶	MNIST مُد	۲.۶۶
۳۱۷	CIFAR-10	۳.۶۶
۳۱۸	ImageNet	۴.۶۶
۳۱۹	COCO میکروسافت	۵.۶۶
۳۱۹	تصاویر باز گوگل	۶.۶۶
۳۲۰	Kaggle	۷.۶۶
۳۲۰	خلاصه فصل	۷.۶
۳۲۲	پروژه ۱: شبکه پیش آموزش دیده به عنوان استخراجگر ویژگی	۸.۶
۳۲۸	پروژه ۲: تدقیق	۹.۶

فصل ۷ / آشکارسازی شیء

۳۳۹	۱.۷ قالب کلی آشکارسازی شیء
۳۴۰	۱.۱.۷ پیشنهاد های ناحیه
۳۴۱	۲.۱.۷ پیش بینی های شبکه
۳۴۲	۳.۱.۷ سرکوب غیربیشینه
۳۴۴	۴.۱.۷ متریک های ارزیابی آشکارساز شیء
۳۴۸	۲.۷ شبکه های کانولوشن ناحیه پایه (RCNN-ها)
۳۴۸	۱.۲.۷ RCNN
۳۵۳	۲.۲.۷ RCNN سریع
۳۵۶	۳.۲.۷ RCNN سریعتر
۳۶۶	۴.۲.۷ جمع بندی خانواده RCNN
۳۶۹	۳.۷ آشکارسازی تک قالب
۳۷۰	۱.۳.۷ معماری سطح کلان SSD
۳۷۳	۲.۳.۷ شبکه پایه
۳۷۵	۳.۳.۷ لایه های ویژگی چندمقیاس
۳۸۰	۴.۳.۷ سرکوب غیربیشینه
۳۸۱	۴.۷ فقط یک بار نگاه (YOLO)
۳۸۳	۱.۴.۷ نحوه کار YOLOv3
۳۸۶	۲.۴.۷ معماری
۳۸۸	۵.۷ خلاصه فصل

فصل ۸ / شبکه های تقابلی مولد

۳۹۳	۱.۸ معماری GAN-ها
۳۹۵	۱.۱.۸ مدل تمایزگر
۳۹۸	۲.۱.۸ مدل مولد
۴۰۱	۳.۱.۸ آموزش GAN
۴۰۳	۱.۳.۱.۸ آموزش تمایزگر
۴۰۳	۲.۳.۱.۸ آموزش ژنراتور (مدل ترکیبی)
۴۰۵	۴.۱.۸ تابع مینی ماکس GAN
۴۰۸	۲.۸ ارزیابی مدل های GAN
۴۱۰	۱.۲.۸ نمره تکوین

۴۱۰	فاصله تکوین فرشیت (FID)	۲.۲.۸
۴۱۱	شیمای استفاده شده در ارزیابی	۳.۲.۸
۴۱۲	کاربردهای مورد اقبال GAN-ها	۳.۸
۴۱۲	سنتز متن به عکس	۱.۳.۸
۴۱۳	ترجمه تصویر به تصویر (Pix2Pix GAN)	۲.۳.۸
۴۱۵	GAN با دقت زیاد	۳.۳.۸
۴۱۵	نمونه‌های اضافی	۴.۳.۸
۴۱۶	GAN پروژه	۴.۸
۴۱۷	کتابخانه‌های ورودی	۱.۴.۸
۴۱۷	دانلود و مصورسازی مجموعه داده	۲.۴.۸
۴۱۸	ساخت مولد	۳.۴.۸
۴۲۰	ساخت تمایزگر	۴.۴.۸
۴۲۱	ساخت مدل ترکیبی	۵.۴.۸
۴۲۲	ساخت تابع آموزش	۶.۴.۸
۴۲۳	نتایج آموزش و مشاهده	۷.۴.۸
۴۲۵	سخن پایانی	۸.۴.۸
۴۲۵	خلاصه فصل	۵.۸

فصل ۹ / DeepDream و انتقال سبک عصبی

۴۲۷	نگاه شبکه‌های کانولوشن به دنیا	۱.۹
۴۲۸	بازبینی نحوه کار شبکه‌های عصبی	۱.۱.۹
۴۲۹	مصورسازی ویژگی‌های CNN	۲.۱.۹
۴۳۰	پیاده‌سازی مصورساز ویژگی	۳.۱.۹
۴۳۴	جمع‌بندی	۴.۱.۹
۴۳۸	DeepDream	۲.۹
۴۳۸	نحوه کار الگوریتم DeepDream	۱.۲.۹
۴۴۰	پیاده‌سازی DeepDream در کراس	۲.۲.۹
۴۴۳	جمع‌بندی	۳.۲.۹
۴۴۸	انتقال سبک عصبی	۳.۹
۴۴۸	اتلاف محتوا	۱.۳.۹
۴۵۱	اتلاف سبک	۲.۳.۹
۴۵۳	اتلاف واریانس کل	۳.۳.۹
۴۵۵	آموزش شبکه	۴.۳.۹
۴۵۷	جمع‌بندی	۵.۳.۹

فصل ۱۰ / نهفته‌سازی بصری

۴۵۹	کاربردهای نهفته‌سازی بصری	۱.۱۰
۴۶۱	تشخیص چهره (FR)	۱.۱.۱۰
۴۶۱	سیستم‌های توصیه‌گر تصویر	۲.۱.۱۰
۴۶۲	بازشناسایی شیئی	۳.۱.۱۰
۴۶۴	یادگیری نهفته‌سازی	۲.۱۰
۴۶۶	توابع اتلاف	۳.۱۰

۴۶۷	برپایش مسئله و ساختارمهندسازی	۱.۳.۱۰
۴۶۸	اتلاف انترویی متقابل	۲.۳.۱۰
۴۷۰	اتلاف کانتراستی	۳.۳.۱۰
۴۷۱	اتلاف سه‌تایی	۴.۳.۱۰
۴۷۴	کاوش داده‌های مفید	۴.۱۰
۴۷۵	بارگذار داده	۱.۴.۱۰
۴۷۸	کاوش داده‌های مفید: یافتن سه‌تایی‌ها	۲.۴.۱۰
۴۷۸	داده‌های نیمه‌سخت	۱.۲.۴.۱۰
۴۸۰	بیج همه (BA)	۳.۴.۱۰
۴۸۱	بیج سخت (BH)	۴.۴.۱۰
۴۸۴	وزندهی بیج (BW)	۵.۴.۱۰
۴۸۴	نمونه بیج (BS)	۶.۴.۱۰
۴۸۵	پروژه: آموزش شبکه نهفته‌سازی	۵.۱۰
۴۸۶	فعالیت ۱: مد-کاوش آیت‌های مشابه	۱.۵.۱۰
۴۸۷	فعالیت ۲: بازشناسایی خودرو	۲.۵.۱۰
۴۸۸	پیاده‌سازی	۳.۵.۱۰
۴۹۰	تست مدل آموزش دیده	۶.۱۰
۴۹۰	فعالیت ۱: بازیابی داخل مغازه	۱.۶.۱۰
۴۹۲	فعالیت ۲: بازشناسایی خودرو	۲.۶.۱۰
۴۹۵	توسعه مرزهای صحت فعلی	۷.۱۰
۴۹۷	خلاصه فصل و نکات	۸.۱۰
۴۹۸	پیوست الف / برپایش	

فصل ۱

مقدمه بینایی کامپیوتر

بینایی کامپیوتر یکی از زمینه‌هایی است که به علت پیشرفت‌های عظیم AI و یادگیری عمیق صورت گرفته در چند سال اخیر، به سرعت پیشرفت کرده است. شبکه‌های عصبی اکنون به اتوموبیل‌های خودران اجازه می‌دهند خودروها و عابران را تشخیص داده، تا با آنها برخورد نکنند. کاربردهای بینایی کامپیوتر در زندگی روزمره، از دوربین‌های امنیتی تا قفل درب‌ها، در وسایل و ابزار هوشمند در حال گسترش است. شناسایی چهره نیز به مراتب بهتر از گذشته شده و مشاهده می‌شود که تلفن‌های همراه می‌توانند چهره را برای باز کردن تلفن و قفل‌های هوشمند برای باز کردن درب، شناسایی نمایند. جای تعجب نخواهد بود که در آینده‌ای نزدیک، میل یا تلویزیون شما بتوانند افراد خاص را در منزل شناسایی کرده و متناسب با علایق و سلیقه شخصی آنها عکس‌العمل نشان دهند. فقط شناسایی اشیاء مطرح نیست؛ یادگیری عمیق کامپیوترها را قادر ساخته موارد جدیدی مثل هنر که قبلاً وجود نداشته، اشیاء جدید، و حتی یکتا، چهره‌های واقع‌گرایانه انسان‌ها، را تصور و ایجاد نمایند.

امکانات نامحدود پژوهش‌های بینایی رایانه و یادگیری ابزار ذی‌ربط، شما را قادر می‌سازد که شاید بتوانید بعضی از این فرآورده‌ها و کاربردهای جدید را خودتان اختراع کنید. حتی اگر در زمینه بینایی کامپیوتر به صورت تخصصی ادامه کار ندهید، بسیاری از مفاهیم ارائه شده در این کتاب برای بعضی از الگوریتم‌ها و معماری‌های یادگیری عمیق مفید خواهد بود. علت این است که، در حالی که تمرکز اصلی بر کاربردهای بینایی کامپیوتر است، مهمترین معماری‌های یادگیری عمیق مثل شبکه‌های عصبی مصنوعی (ANN)، شبکه‌های کانولوشن (CNN-ها)، شبکه‌های برگشتی (RNN-ها)، و بسیاری دیگر نیز پوشش داده شده، که قابل انتقال به زمینه‌های دیگر نظیر