گزارش تحلیلی معماری IAFMNet در بازسازی فوق‌دقیق تصاویر

(SISR)

درس: پردازش تصویر دیجیتال

استاد: دکتر غیاثی راد

تابستان ۱۴۰۵

محبوبه کریمی احمد آبادی

۴۰۴۱۲۶۳۴

فهرست مطالب

تبیین مسئله ۲

آناتومی سیستم و اکوسیستم داده ۲

جدول مشخصات اکوسیستم داده‌ها ۳

تشریح متدولوژی: نوآوری در تعدیل ویژگی‌های آگاه از اطلاعات ۴

اجزای کلیدی معماری ۴

ارزیابی عملکرد و تحلیل نتایج تجربی ۵

یافته‌های کلیدی ارزیابی ۵

محدودیت‌ها، سنتز نهایی و افق‌های توسعه ۶

نقشه راه توسعه و پیشنهادات فنی ۶

نتیجه‌گیری نهایی ۶

تبیین مسئله

در پارادایم کنونی بازسازی فوق‌دقیق تصاویر (SISR) ، اکثر مدل‌های پیشرفته (اعم از مبتنی بر CNN یا ترنسفورمر) از یک رویکرد «یکنواخت فضایی» (Spatially Uniform) پیروی می‌کنند؛ بدین معنا که منابع محاسباتی یکسانی را برای تمامی پیکسل‌ها، صرف‌نظر از میزان پیچیدگی ساختاری آن‌ها، تخصیص می‌دهند. از منظر تئوری اطلاعات شانون و مدل‌سازی «نرخ-اعوجاج» (Rate-Distortion) این استراتژی به شدت ناکارآمد است.

در فرآیند تخریب تصویر، لبه‌ها و بافت‌های فرکانس‌بالا اطلاعات بیشتری از دست می‌دهند و بازسازی آن‌ها به «نرخ بیت» یا ظرفیت محاسباتی بالاتری نیاز دارد، در حالی که نواحی صاف (مانند آسمان) با هزینه اندک قابل بازیابی هستند. تغییر نگرش از پردازش پیکسل‌محورِ ایستا به سمت **«تخصیص پویای نرخ بیت» (Dynamic Rate Allocation)** بر پایه چگالی اطلاعات، یک جهش استراتژیک در بهره‌وری سیستم‌های بینایی ماشین محسوب می‌شود. در حالی که روش‌های سنتی ریاضی مانند KNN یا Bilinear صرفاً با بازآرایی داده‌های موجود عمل می‌کنند، رویکرد **IAFMNet** بر شناسایی هوشمند مناطقی تمرکز دارد که بازسازی آن‌ها بیشترین چالش را دارد. این مدل با تمرکز بر ۵٪ از پیکسل‌های حیاتی، توازن میان وفاداری بصری و سربار محاسباتی (FLOPs) را به نفع سیستم‌های بلادرنگ تغییر می‌دهد.

آناتومی سیستم و اکوسیستم داده

در توسعه مدل‌های مولد، کیفیت داده‌های مرجع تعیین‌کننده سقف یادگیری مدل است. در سناریوی بازسازی x4، بعاد تصویر در هر ضلع ۴ برابر بزرگتر می‌شود که منجر به افزایش ۱۶ برابری در تعداد پیکسل‌ها و پیچیدگی نمایی محاسبات می‌گردد. مدیریت این حجم از داده‌ها در IAFMNet از طریق بهینه‌سازی FLOPs محاسبه شده بر مبنای خروجی استاندارد 720p انجام می‌شود تا از کارایی مدل در مقیاس عملیاتی اطمینان حاصل گردد.

مدل IAFMNet تصویر LR (کم‌کیفیت) را به عنوان ورودی دریافت کرده و علاوه بر تصویرSR ، یک «نقشه چگالی اطلاعات» (IDM) تولید می‌کند که بیانگر هزینه کدگذاری (Encoding Cost) در نواحی مختلف است.

جدول مشخصات اکوسیستم داده‌ها

نوع داده	منبع(Source)	اندازه / جزئیات	تکنیک‌های پیش‌پردازش و تقویت
مجموعه‌داده آموزشی	**DF2K** ترکیب (Flickr2KوDIV2K)	۳۴۵۰ تصویر با کیفیت بالا (HR)	استفاده از **Bicubic Downsampling** برای تولید تصاویر کم‌کیفیت(LR)
پچ‌های آموزشی	استخراج شده از مجموعه‌داده **DF2K**	برش‌های **64x64** پیکسل	**Flip**(وارونگی) تصادفی و **چرخش ۹۰ درجه**
داده‌های ارزیابی(Test)	**Urban100, Set5, Manga109**, Set14, BSD100	ابعاد متنوع تا خروجی ۷۲۰p	محاسبه معیارهای **PSNR/SSIM** بر روی **کانال Y**(روشنایی)
خروجی سیستم	بازسازی تصویر و تخمین اطلاعات	مقیاس بزرگ‌نمایی **x2, x3, x4**	**Residual Learning** ترکیب خروجی شبکه با تصویر(Bicubic)

تشریح متدولوژی: نوآوری در تعدیل ویژگی‌های آگاه از اطلاعات

طراحی IAFMNet عمیقاً با مفاهیم تئوری اطلاعات گره خورده است. قلب این معماری، تابع هزینه **Information Entropy Loss (L_{IE})** است که به صورت کاملاً بدون نظارت مدل را برای شناسایی نواحی با چگالی اطلاعات بالا هدایت می‌کند.

اجزای کلیدی معماری

- **تخمین‌گر چگالی اطلاعات (IDE)** : این واحد از لایه‌های **GDN (Generalized Divisive Normalization)** برای مدل‌سازی دقیق چگالی استفاده می‌کند. IDE ویژگی‌ها را به دو مولفه میانگین (μ) و مقیاس(θ) یا همان IDM تبدیل می‌کند. در زمان آموزش، با تزریق **نویز یکنواخت** به ویژگی‌ها، فرآیند کوانتیزاسیون مشتق‌پذیر شده و امکان تخمین انتروپی فراهم می‌گردد.
- **تخصیص منابع هدایت‌شده - (IGRA) شاخه سخت** : این شاخه از **Submanifold Sparse Convolution (SSC)** بهره می‌برد. برخلاف کانولوشن‌های اسپارس معمولی SSC ، الگوی اسپارسیته را در تمامی لایه‌ها حفظ کرده و تنها ۵٪ از پیکسل‌های کلیدی (Top-k) را پردازش می‌کند که منجر به صرفه‌جویی ۹۵ درصدی در محاسبات این شاخه می‌شود.
- **ماژول بازتنظیم افاین - (ARM) شاخه نرم:** این بخش به تقویت تطبیقی تمامی ویژگی‌ها می‌پردازد، جایی که شدت تعدیل افاین مستقیماً توسط مقادیر پیوسته IDM کنترل می‌شود.
- **منطق فرآیند عملیاتی (Operational Logic)**

۱. **استخراج ویژگی:** تبدیل تصویر ورودی به ویژگی‌های ۲۴ کاناله اولیه.

۲. **تخمین چگالی:** تولید نقشه IDM توسط IDE برای شناسایی اولویت‌های اطلاعاتی.

۳. **توزیع کانال (Split)** : تقسیم ویژگی‌ها به دو بخش برای پردازش در دو شاخه موازی.

۴. **پردازش دوگانه:** پردازش همزمان در IGRA (تمرکز بر لبه‌های سخت) و ARM (تعدیل نرم بافت‌ها).

۵. **بازسازی** : ادغام ویژگی‌های تقویت شده و استفاده از PixelShuffle برای تولید تصویر نهایی.

ارزیابی عملکرد و تحلیل نتایج تجربی

برای سنجش وفاداری بازسازی، از معیارهای PSNR (شفافیت) و SSIM (شباهت ساختاری) استفاده شده است. تحلیل‌های معماری نشان می‌دهد که IAFMNet موفق شده است با پارامترهای بسیار کمتر، عملکردی در سطح یا فراتر از مدل‌های سنگین ارائه دهد.

یافته‌های کلیدی ارزیابی

- **بهره‌وری پارامتریک** : نسخه اصلی **IAFMNet (220K Params)** و نسخه سبک **IAFMNet-S (166K Params)** در مقایسه با مدل‌هایی نظیر SwinIR-light (930K) یا SRFormer-light (873K)، PSNR بالاتری را با کسری از پیچیدگی محاسباتی (FLOPs) ثبت کرده‌اند (به عنوان مثال ۳۲.۳۷ در Set5).
- **تحلیل نرخ بهبود** : در پیاده‌سازی‌های اولیه (Steps 2000) ، بهبود ۹۷.۸ درصدی در L_{sr} مشاهده شد. با این حال، در شرایط SOTA و با ۱ میلیون گام آموزشی، مدل به پایداری خیره‌کننده‌ای در بازسازی الگوهای موازی Urban100 دست یافته است.
- **برتری نسبت به اپراتورهای سنتی** : برخلاف فیلترهای Sobel و Laplacian که صرفاً بر گرادیان‌های شدید تمرکز دارند ، IDM به دلیل ماهیت انتروپی‌محور، قادر به شناسایی بافت‌های **کم‌کنتراست** (Low-contrast) مانند خز حیوانات یا بافت‌های ظریف پس‌زمینه است که فیلترهای کلاسیک در شناسایی آن‌ها شکست می‌خورند.

محدودیت‌ها، سنتز نهایی و افق‌های توسعه

معماری IAFMNet ثابت کرد که در عصر محدودیت‌های سخت‌افزاری، «پردازش هوشمند» بر «پردازش سنگین» ارجحیت دارد. استفاده از Sparse Convolution بار محاسباتی را دقیقاً بر نقاط بحرانی متمرکز می‌کند که این ویژگی IAFMNet ، را به گزینه‌ای ایده‌آل برای **Edge Devices** و گوشی‌های هوشمند تبدیل کرده است.

نقشه راه توسعه و پیشنهادات فنی

- **بهینه‌سازی پارامتریک** : بر اساس مطالعات Ablation، تنظیم دقیق ضریب **L_{IE} = 0.0001** برای ایجاد تعادل بهینه میان دقت بازسازی و تخمین انتروپی حیاتی است.
- **توسعه عملیاتی** : انتقال کامل فرآیند آموزش به محیط GPU و استفاده از Learning Rate Scheduler برای عبور از چالش Overfitting در گام‌های بالا.
- **ارتقای معماری** : پیاده‌سازی کامل ۶ بلوک IFEB در نسخه تجاری برای دستیابی به حداکثر وفاداری بصری.

نتیجه‌گیری نهایی

IAFMNet با پیوند زدن تئوری اطلاعات به مهندسی شبکه‌های عصبی، استاندارد جدیدی در SISR سبک‌وزن تعریف کرده است. این مدل با جایگزینی استراتژی‌های یکنواخت با مکانیسم‌های آگاه از اطلاعات، راهکاری مقیاس‌پذیر و هوشمند برای چالش بازسازی تصاویر در دنیای دیجیتال فراهم آورده است.