



بررسی تکنیک‌ها و پیشرفت‌های نوین ادغام تصویر

فاطمه جوکار

مربی، گروه مهندسی کامپیوتر، دانشکده مهندسی برق و کامپیوتر، دانشگاه فنی و حرفه ای شهید باهنر شیراز

محمد رضا بحرانی

مربی، گروه مهندسی کامپیوتر، دانشکده مهندسی برق و کامپیوتر، دانشگاه فنی و حرفه ای شهید باهنر شیراز

سید آرش موسوی

دانشجوی کارشناسی، گروه مهندسی کامپیوتر، دانشگاه امیرکبیر، تهران

چکیده

ادغام تصویر ادغام تصویر فرآیندی است که داده‌های چندین تصویر از یک صحنه را ترکیب کرده و تصویری جامع‌تر و با کیفیت بالاتر ایجاد می‌کند. این تکنیک در زمینه‌هایی مانند سنجش از دور، تصویربرداری پزشکی و بینایی کامپیوتری کاربردهای فراوانی دارد. در ابتدا، روش‌های سنتی مانند میانگین‌گیری وزنی و فیلتر لاپلاس برای ترکیب تصاویر استفاده می‌شدند، اما با پیشرفت‌های فناوری، روش‌های مبتنی بر یادگیری عمیق مانند شبکه‌های عصبی پیچشی (CNN) و شبکه‌های مولد تخاصمی (GAN) به کار گرفته شده‌اند که دقت و کیفیت ادغام تصویر را به طرز چشمگیری بهبود می‌بخشند.

این پیشرفت‌ها توانسته‌اند بسیاری از چالش‌ها مانند حفظ جزئیات تصویر، کاهش نویز، و بهینه‌سازی پردازش را حل کنند. به ویژه، الگوریتم‌های جدید به کاهش پیچیدگی محاسباتی و افزایش دقت در ادغام تصاویر کمک کرده‌اند. علاوه بر این، این تکنیک‌ها در کاربردهای گسترده‌ای از جمله تشخیص بیماری‌ها در تصویربرداری پزشکی، بهبود کیفیت تصاویر ماهواره‌ای در سنجش از دور و افزایش دقت در تشخیص اشیا در بینایی کامپیوتری مورد استفاده قرار می‌گیرند. این مقاله به بررسی روش‌های مختلف ادغام تصویر، چالش‌ها و پیشرفت‌های اخیر در این حوزه پرداخته است.

واژگان کلیدی: ادغام تصویر، تصویربرداری پزشکی، سنجش از دور، همجوشی تصویر، کیفیت تصویر.

مقدمه

محدودیت‌های سخت‌افزاری و شرایط محیطی معمولاً با کیفیت کمتری ثبت می‌شوند. ادغام تصویر با ترکیب داده‌های حاصل از سنسورهای مختلف، وضوح و کیفیت این تصاویر را به طرز چشمگیری بهبود می‌بخشد. در تصویربرداری پزشکی، ادغام تصاویر حاصل از روش‌های مختلف مانند MRI، CT و PET می‌تواند به تشخیص دقیق‌تر بیماری‌ها و برنامه‌ریزی درمان کمک کند. همچنین در بینایی کامپیوتری، ادغام تصویر به افزایش دقت در تشخیص اشیاء و ردیابی آن‌ها کمک می‌کند و به ویژه در کاربردهای نظامی و امنیتی از اهمیت بالایی برخوردار است.

برخوردار است. با این حال، ادغام تصویر با چالش‌ها و محدودیت‌هایی نیز مواجه است. حفظ لبه‌ها و جزئیات مهم در فرآیند ترکیب، تفاوت‌های روشنایی و کنتراست میان تصاویر مختلف و وجود نویز و سیگنال‌های ناخواسته از جمله مشکلاتی هستند که باید به دقت مدیریت شوند. علاوه بر این، پیچیدگی‌های محاسباتی و نیاز به منابع بالا در برخی از روش‌های مدرن ادغام تصویر، می‌تواند بر کارایی این تکنیک‌ها تأثیر بگذارد. به رغم این چالش‌ها، پیشرفت‌های اخیر در یادگیری عمیق و تکنیک‌های نوین ادغام تصویر نویدبخش آینده‌ای روشن برای این حوزه است. با توجه به افزایش قدرت محاسباتی و دسترسی به داده‌های بزرگتر، انتظار می‌رود که روش‌های جدید ادغام تصویر به طور گسترده‌تری در دستگاه‌های مختلف و در زمینه‌های مختلف علمی و صنعتی قابل استفاده باشند. به علاوه، کاربردهای نوین در زمینه‌های واقعیت افزوده و مجازی، پزشکی دقیق و نظارت هوشمند نیز از قابلیت‌های بالقوه ادغام تصویر بهره‌مند خواهند شد.

مفهوم ادغام تصویر

ادغام تصویر فرآیند ترکیب اطلاعات چندین تصویر برای به دست آوردن تصویر نهایی با دقت و کیفیت بالا است. ادغام تصویر به معنای ترکیب داده‌های چندین تصویر از یک صحنه به گونه ای است که تصویر نهایی اطلاعات کامل تری نسبت به تصاویر اولیه ارائه دهد. به طور کلی، این فرآیند می‌تواند با استفاده از فرمول‌های ریاضی زیر مدلسازی شود:

$$(1) \quad I_{final}(x,y) = f(I_1(x,y), I_2(x,y), \dots, I_n(x,y))$$

در اینجا، $I_{final}(x,y)$ تصویر نهایی، f تابع ترکیب و $I_1(x,y), I_2(x,y), \dots, I_n(x,y)$ تصاویر ورودی هستند.

ادغام تصویر از دهه ۱۹۷۰ با استفاده از روش‌های ابتدایی آغاز شد و با پیشرفت فناوری به یک حوزه پیچیده و پیشرفته تبدیل گردید. در دهه‌های اخیر، روش‌های یادگیری عمیق مانند شبکه‌های عصبی پیچشی (CNN) و شبکه‌های مولد تخصصی (GAN) به طور قابل توجهی توانایی ادغام تصویر را بهبود بخشیده‌اند. این تکنیک به دلیل توانایی‌اش در بهبود کیفیت تصویر و استخراج اطلاعات دقیق‌تر از منابع مختلف، به طور گسترده‌ای در حوزه‌های مختلف علمی و صنعتی مورد استفاده قرار می‌گیرد.

در سنجش از دور، ادغام تصاویر می‌تواند اطلاعات به دست آمده از حسگرهای مختلف مانند حسگرهای نوری، راداری و حرارتی را ترکیب کند. این ترکیب به محققان این امکان را می‌دهد که ویژگی‌های زمین‌شناختی و زیست‌محیطی را با دقت بیشتری تحلیل کنند و به طور موثرتری تغییرات محیطی را شناسایی کنند. به عنوان مثال، ادغام تصاویر ماهواره‌ای با داده‌های راداری می‌تواند در شناسایی مناطق آسیب‌دیده در برابر بلایای طبیعی و ارزیابی تأثیرات آن‌ها بسیار مفید باشد.

در تصویربرداری پزشکی، ادغام تصاویر به پزشکان این امکان را می‌دهد که اطلاعات حاصل از تکنیک‌های تصویربرداری مختلف مانند CT، MRI و PET را با هم ترکیب کنند. این ادغام می‌تواند به تشخیص بهتر بیماری‌ها و تحلیل دقیق‌تر وضعیت بیمار کمک کند. به عنوان مثال، ترکیب تصاویر CT و MRI می‌تواند اطلاعات ساختاری و عملکردی را به طور همزمان ارائه دهد و در نتیجه تشخیص و درمان بهتری را فراهم آورد.

در زمینه بینایی کامپیوتری، ادغام تصاویر می‌تواند در بهبود قابلیت‌های شناسایی و تحلیل اشیاء کمک کند. با ترکیب داده‌های تصاویر مختلف، سیستم‌های بینایی کامپیوتری می‌توانند اطلاعات بیشتری درباره محیط اطراف خود جمع‌آوری کنند و در نتیجه دقت و کارایی بیشتری در وظایف مختلف، از جمله شناسایی چهره، تشخیص اشیاء و ناوبری، داشته باشند.

ادغام تصویر به فرآیند ترکیب اطلاعات چندین تصویر به گونه‌ای گفته می‌شود که تصویر نهایی ویژگی‌های غنی‌تری نسبت به تصاویر اولیه داشته باشد. این تکنیک به سه دسته اصلی ادغام پیکسل‌محور، ادغام ویژگی‌محور و ادغام تصمیم‌محور تقسیم می‌شود. هر کدام از این روش‌ها بر اساس نوع داده‌ها و اهداف خاص متفاوتی به کار می‌روند.

ادغام پیکسل‌محور

این روش بر اساس ترکیب اطلاعات پیکسل‌های تصاویر ورودی عمل می‌کند. ادغام پیکسل‌محور به استفاده از مقادیر پیکسل‌های هر تصویر و ترکیب آن‌ها برای تولید تصویر نهایی می‌پردازد. این نوع ادغام برای تصاویری که نیاز به دقت پیکسل به پیکسل دارند، مناسب است. میانگین‌گیری، وزن‌دهی و فیلترهای لاپلاسی نمونه‌های رایج این روش هستند.

- میانگین‌گیری

$$(2) \quad \frac{I_1(x,y) + I_2(x,y) + \dots + I_n(x,y)}{n} = I_{merged}(x,y)$$

یکی از ساده‌ترین روش‌ها برای ادغام تصاویر، محاسبه میانگین مقادیر پیکسل‌ها در هر موقعیت است.

در اینجا، $I_n(x,y)$ مقدار روشنایی پیکسل در مختصات (x,y) تصویر n ام است، $I_{merged}(x,y)$ مقدار روشنایی پیکسل ادغام شده در مختصات (x,y) است که n در آن تعداد تصاویر است.

- وزن‌دهی

در این روش، به هر تصویر وزنی اختصاص داده می‌شود و سپس مقادیر پیکسل‌ها به طور وزنی ترکیب می‌شوند.

$$(3) \quad I_1(x,y) \cdot w_1 + I_2(x,y) \cdot w_2 + \dots + I_n(x,y) \cdot w_n = I_{merged}(x,y)$$

در اینجا، n تعداد تصاویر است، w_1, w_2, \dots, w_n وزن‌های اختصاص یافته به هر تصویر هستند و مجموع وزن‌ها باید برابر با ۱ باشد یعنی $w_1 + w_2 + \dots + w_n = 1$.

- فیلترهای لاپلاسی

این روش شامل ایجاد هرم لاپلاسی از تصاویر است که در آن اطلاعات لبه‌ها و جزئیات هر تصویر حفظ می‌شود. سپس لایه‌های مختلف هرم ترکیب می‌شوند تا تصویر نهایی ایجاد شود. فیلتر لاپلاس معمولاً با یک ماسک مانند زیر اعمال می‌شود.

$$L = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix}$$

هر پیکسل (x,y) در تصویر، پس از اعمال این ماسک، مقدار جدیدی می‌گیرد که لبه‌ها و جزئیات تصویر را برجسته می‌کند:

$$(4) \quad L * I_i(x,y) = \text{Laplace}(I_i)(x,y)$$

در اینجا، $*$ نشان دهنده عملیات کانولوشن بین تصویر I_i و ماسک L است. پس از استخراج لبه‌ها و جزئیات با فیلتر لاپلاس، می‌توانیم این جزئیات را به تصاویر اصلی اضافه کنیم. به این صورت که تصویر لاپلاسی هر تصویر را با وزن α به تصویر اصلی اضافه کنیم.

$$(5) \quad \text{Laplace}(I_i)(x,y) \cdot \alpha + I_i(x,y) = I_{enhanced}(x,y)$$

مقدار α ضریبی است که شدت اثر فیلتر را تنظیم می کند و معمولاً مقداری بین 0.2 تا 0.5 انتخاب می شود.

ادغام ویژگی محور

ادغام ویژگی محور بر اساس استخراج و ترکیب ویژگی های مهم از تصاویر مختلف انجام می شود. در این روش، ابتدا ویژگی ها از تصاویر ورودی استخراج می شوند و سپس این ویژگی ها ترکیب می شوند. تحلیل مؤلفه های اصلی (PCA)، ایجاد ویژگی های سطحی و روش های یادگیری عمیق رایج ترین نمونه های این روش هستند.

- تحلیل مؤلفه های اصلی (PCA)

این روش برای کاهش ابعاد داده ها و استخراج ویژگی های اصلی استفاده می شود. ابتدا ماتریس ویژگی های تصاویر را تشکیل می دهیم. فرض کنیم X ماتریس ویژگی ها باشد که هر ستون آن، ویژگی های یک تصویر است. اگر تعداد ویژگی ها n و تعداد تصاویر m باشد، X ابعاد $m \times n$ دارد. سپس هر ویژگی را از میانگین آن کم می کنیم تا ماتریس میانگین گیری شده $X_{centered}$ به دست آید.

$$(6) \quad X_{centered} = X - \bar{X}$$

که \bar{X} میانگین هر ویژگی است. سپس ماتریس کوواریانس C از ماتریس میانگین گیری شده به صورت زیر به دست می آید:

$$(7) \quad C = \frac{1}{m-1} X_{centered} \cdot X_{centered}^T$$

سپس مقادیر و بردارهای ویژه ماتریس کوواریانس C را پیدا کنید. این مقادیر ویژه نشان دهنده اهمیت هر مؤلفه اصلی هستند. با مرتب کردن بردارهای ویژه بر اساس مقادیر ویژه متناظر، مؤلفه های اصلی انتخاب شده را به عنوان ماتریس انتقال W استفاده کنید. ماتریس ویژگی های اصلی یا ادغام شده را به صورت زیر محاسبه کنید:

$$(8) \quad X_{pca} = W^T \cdot X_{centered}$$

در نهایت ماتریس به دست آمده شامل ویژگی های ادغام شده تصاویر است.

ایجاد ویژگی های سطحی: در این روش، ویژگی هایی مانند لبه ها، بافت ها و رنگ ها از تصاویر استخراج می شود.

روش های یادگیری عمیق: شبکه های عصبی مانند CNN می توانند به طور خودکار ویژگی های پیچیده از تصاویر استخراج کنند.

ادغام تصمیم محور

این نوع ادغام به ترکیب نتایج حاصل از چندین سیستم شناسایی یا تشخیص پرداخته و تصمیم نهایی را براساس خروجی های آن ها می سازد.

- رای گیری

در این روش، هر سیستم شناسایی یک رأی به خروجی خود می دهد و خروجی نهایی بر اساس رأی اکثریت تعیین می شود. رای گیری و ترکیب خطی دو نمونه رایج این روش هستند.

$$(9) \quad C_{final} = \argmax(\sum_{i=1}^N \delta(C_{i,c}))$$

که در آن $\delta(C_{i,c})$ یک تابع است که ۱ را به ازای تطابق طبقه C_i با طبقه c برمی گرداند. در غیر این صورت صفر است.

- ترکیب خطی

این روش به ترکیب خروجی های سیستم های مختلف بر اساس وزن های از پیش تعیین شده می پردازد.

$$(10) \quad C_{\text{final}} = \operatorname{argmax}(\sum_{i=1}^N w_i \cdot \delta(C_{i,c}))$$

چالش‌ها و محدودیت‌های ادغام تصویر

ادغام تصویر یک فرآیند پیچیده است که با وجود پیشرفت‌های قابل توجه در تکنیک‌ها و فناوری‌های مرتبط، همچنان با چالش‌ها و محدودیت‌های متعددی مواجه است. این چالش‌ها می‌توانند تأثیر مستقیمی بر کیفیت تصویر نهایی و دقت اطلاعات استخراج‌شده داشته باشند. در ادامه به بررسی برخی از این چالش‌ها و محدودیت‌ها خواهیم پرداخت که در کاربردهای مختلف ادغام تصویر، از جمله سنجش از دور و تصویربرداری پزشکی، نمود بیشتری دارند (جدول ۱).

– حفظ لبه‌ها و جزئیات مهم

یکی از بزرگ‌ترین چالش‌ها در ادغام تصویر، حفظ لبه‌ها و جزئیات حیاتی در فرآیند ترکیب است. در بسیاری از روش‌ها، لبه‌های تصویر ممکن است در نتیجه ترکیب محو شوند یا نادیده گرفته شوند، که این امر می‌تواند به کیفیت و دقت تصویر نهایی آسیب برساند. این موضوع به‌ویژه در کاربردهایی مانند تصویربرداری پزشکی که در آن تشخیص دقیق ویژگی‌ها و ساختارها اهمیت دارد، حائز اهمیت است. تفاوت روشنایی و کنتراست: تصاویر گرفته‌شده از سنسورهای مختلف ممکن است دارای سطوح متفاوتی از روشنایی و کنتراست باشند. این تفاوت‌ها می‌تواند در فرآیند ادغام مشکلاتی ایجاد کند، زیرا ترکیب تصاویر با روشنایی‌های مختلف ممکن است منجر به عدم هماهنگی در تصویر نهایی شود. به همین دلیل، نیاز به تنظیمات دقیق و پیش‌پردازش برای همسان‌سازی روشنایی و کنتراست وجود دارد.

نویز و سیگنال‌های ناخواسته: وجود نویز در تصاویر یکی دیگر از مسائل رایج در ادغام تصویر است. نویز می‌تواند جزئیات مهم تصویر را کاهش دهد و به‌طور کلی کیفیت تصویر نهایی را تحت تأثیر قرار دهد. اگرچه روش‌های مختلفی برای حذف نویز وجود دارد، اما حفظ جزئیات و دقت تصویر نهایی به‌ویژه در تصاویر پزشکی یا تصاویر مربوط به سنجش از دور یک چالش مهم باقی می‌ماند.

پیچیدگی محاسباتی: روش‌های پیچیده‌تری مانند شبکه‌های عصبی عمیق (DNNs) نیازمند توان محاسباتی بالا و دسترسی به منابع بزرگ داده هستند. این امر ممکن است در سیستم‌های با محدودیت منابع، مانند دستگاه‌های همراه یا سیستم‌های embedded، باعث کاهش کارایی و دشواری در پیاده‌سازی این تکنیک‌ها شود. به همین دلیل، بهینه‌سازی منابع و کاهش پیچیدگی محاسباتی یک چالش عمده در این حوزه است.

انتخاب مناسب تکنیک ادغام: انتخاب روش مناسب برای ادغام تصاویر بسته به نوع و کیفیت تصاویر، یک چالش جدی است. عدم شناخت کامل از ویژگی‌های تصاویر و تکنیک‌های ادغام می‌تواند منجر به انتخاب نادرست و در نتیجه کاهش کیفیت تصویر نهایی شود. برای غلبه بر این چالش، پژوهشگران باید به‌طور دقیق با ویژگی‌های هر تکنیک و نیازهای خاص داده‌های خود آشنا شوند.

جدول ۱: جدول مقایسه‌ای چالش‌ها، روش‌های مقابله و کاربردهای ادغام تصویر

چالش‌ها	روش‌های موجود برای مقابله	کاربردهای مربوطه
حفظ لبه‌ها و جزئیات مهم	استفاده از فیلترهای لاپلاسی و گاوسی	تصویربرداری پزشکی، بینایی کامپیوتری
تفاوت روشنایی و کنتراست	تکنیک‌های تنظیم روشنایی و کنتراست	سنجش از دور، تصویربرداری پزشکی
نویز و سیگنال‌های ناخواسته	فیلترهای میانگین و میانگین وزنی استفاده از فیلترهای دوطرفه و میان‌گذری	تصویربرداری پزشکی، شناسایی شی

پنجیدگی محاسباتی	کاهش ابعاد داده‌ها با تحلیل مؤلفه‌های اصلی (PCA) بهینه‌سازی الگوریتم‌ها و استفاده از شبکه‌های سبک‌تر	بینایی کامپیوتری، سامانه‌های زمان واقعی
انتخاب مناسب تکنیک ادغام	روش‌های بهبود یافته مانند GAN ها و شبکه‌های ترکیبی پیش‌بینی	ادغام تصاویر ماهواره‌ای، تشخیص بیماری‌ها
کاهش دقت در ترکیب	استفاده از الگوریتم‌های بهینه شده و سریع‌تر برای پردازش بهره‌گیری از واحدهای پردازش موازی مانند GPU	شبیه‌سازی، واقعیت افزوده، واقعیت مجازی

محدودیت‌های زمانی: زمان پردازش طولانی می‌تواند در کاربردهای زمان واقعی مانند بینایی کامپیوتری و نظارت بر امنیت یک مشکل باشد. در چنین مواردی، نیاز به پردازش سریع تصاویر و تولید نتایج در زمان واقعی وجود دارد. بنابراین، تکنیک‌های ادغام باید به گونه‌ای طراحی شوند که سرعت عمل را در کنار کیفیت تصویر حفظ کنند.

کاهش دقت در ترکیب: برخی از روش‌های ادغام ممکن است دقت تصاویر نهایی را تحت تأثیر قرار دهند و منجر به ایجاد نتایج غیرواقعی شوند. به عنوان مثال، ترکیب تصاویر با استفاده از روش‌های ساده ممکن است به اطلاعات نادرست یا تشخیص نادرست ویژگی‌ها منجر شود. برای حل این مشکل، به کارگیری روش‌های پیشرفته و یادگیری عمیق می‌تواند کمک‌کننده باشد، اما این امر به افزایش پیچیدگی و نیاز به منابع بیشتر منجر می‌شود.

الگوریتم‌های نوین در ادغام تصویر

در سال‌های اخیر، تکنیک‌های ادغام تصویر در حوزه‌های مختلف مانند سنجش از دور، تصویربرداری پزشکی و بینایی کامپیوتری پیشرفت‌های چشمگیری داشته‌اند. این تکنیک‌ها با هدف ترکیب اطلاعات از تصاویر مختلف برای دستیابی به تصویری با کیفیت بالاتر و جزئیات بیشتر توسعه یافته‌اند. یکی از روش‌های مهم در این زمینه، روش ادغام تصاویر مادون قرمز و مرئی است که در آن با استفاده از ترکیب فیلترهای گاوسی و دوطرفه، اطلاعات هر دو نوع تصویر را در یک تصویر واحد جمع می‌کنند. این روش به دلیل توانایی در حفظ لبه‌ها و کاهش نویز، در زمینه‌هایی همچون ناوبری و تشخیص هدف کارایی بالایی از خود نشان داده است.

الگوریتم‌های مبتنی بر شبکه‌های عصبی پیش‌بینی (CNN) نیز در ادغام تصویر نقش مهمی ایفا کرده‌اند. به عنوان مثال، IFCNN که یک نوع پیشرفته CNN است، با استفاده از تکنیک‌های کانولوشن و به کارگیری ResNet، دقت و سرعت بالایی در آموزش و پردازش تصاویر به دست آورده است. این مدل‌ها با ترکیب ویژگی‌های استخراج شده از تصاویر، جزئیات بیشتری ارائه می‌دهند و در حوزه‌هایی مانند تصویربرداری پزشکی و پردازش تصاویر ماهواره‌ای به کار گرفته می‌شوند.

از دیگر روش‌های نوین، شبکه‌های GAN مبتنی بر ادغام چند مقیاسی هستند. مدل‌هایی مانند MEF-GAN که از تکنیک GAN برای بهبود کیفیت و ترکیب اطلاعات تصاویر با رزولوشن‌های متعدد استفاده می‌کنند، در تولید تصاویر با کیفیت بالا برای کاربردهایی همچون واقعیت مجازی (VR) و واقعیت افزوده (AR) موثر بوده‌اند. در این روش‌ها، شبکه‌های تولیدکننده و تشخیص‌دهنده در کنار هم کار می‌کنند تا تصویر نهایی از نظر بصری واقعی‌تر و باکیفیت‌تر باشد.

مدل ResNetFusion نیز با استفاده از شبکه‌های عصبی عمیق و کانولوشن‌های مبتنی بر ResNet، ویژگی‌های مهم تصویر را حفظ کرده و گرادین‌ها را به صورت موثرتری از لایه‌ها عبور می‌دهد تا تصویری دقیق و با وضوح بالا ارائه دهد. این روش برای کاربردهایی چون رانندگی خودکار و نظارت و تصویربرداری پزشکی بسیار مناسب است.

روش‌های کانولوشنال همچون ConvSR نیز با تمرکز بر ویژگی‌های ساختاری و محلی تصاویر، توانسته‌اند اطلاعات مهم تصاویر را استخراج کرده و به طور موثرتری با یکدیگر ترکیب کنند. این تکنیک‌ها به دلیل وضوح فضایی و طیفی بالا، در تصویربرداری پزشکی کاربرد فراوان دارند و توانایی تشخیص دقیق‌تر نواحی خاص را فراهم می‌کنند.

در نهایت، روش‌های نوینی همچون PanGAN و CBF به ترتیب برای تولید تصاویر پانوراما و حفظ لبه‌های تصویر در فرآیند ادغام معرفی شده‌اند. روش PanGAN با استفاده از GAN، تصاویر گسترده‌ای ایجاد می‌کند که در حوزه‌های AR و VR کاربردی است، و روش CBF با در نظر گرفتن نزدیکی هندسی و شدت پیکسل‌ها، به‌طور ویژه برای تصاویر پزشکی با دقت بالا مناسب است.

به طور کلی، پیشرفت در مدل‌های مبتنی بر یادگیری عمیق و ترکیب آن‌ها با فیلترهای سنتی منجر به توسعه روش‌های ادغام تصاویر با دقت و کیفیت بالاتری شده است که پاسخگوی نیازهای تخصصی در زمینه‌های مختلف تصویربرداری است.

معماری‌های رایج ادغام تصاویر

برای دستیابی به ادغام تصاویر، معماری‌های مختلفی توسعه یافته‌اند. در ادامه برخی از آنها را به طور کلی بررسی می‌کنیم.

- معماری‌های مبتنی بر تبدیل

در این روش‌ها، تصاویر ورودی به حوزه‌های فرکانسی یا مقیاسی مختلف تبدیل می‌شوند و سپس با ترکیب ضرایب تبدیل، تصویر نهایی بازسازی می‌شود.

- معماری‌های مبتنی بر آمار

این روش‌ها بر اساس مدل‌های آماری و احتمالاتی عمل می‌کنند. به عنوان مثال، روش‌های مبتنی بر فیلتر کالمن یا فیلتر بیزین با ترکیب اطلاعات از منابع مختلف، تصویری با دقت بالاتر تولید می‌کنند.

- معماری‌های مبتنی بر یادگیری عمیق

با استخراج ویژگی‌های چندمقیاسی از تصاویر ورودی، ادغام مؤثری را انجام می‌دهند.

- معماری U-Net

این معماری که ابتدا برای بخش‌بندی تصاویر پزشکی توسعه یافت، به دلیل ساختار U شکل خود، در ادغام تصاویر نیز کاربرد دارد. U-Net با استفاده از مسیرهای کاهشی و افزایشی، اطلاعات مکانی و بافتی تصاویر را حفظ کرده و ادغام مؤثری را ارائه می‌دهد.

انتخاب معماری مناسب برای ادغام تصاویر به نوع تصاویر ورودی، کاربرد مورد نظر و منابع محاسباتی در دسترس بستگی دارد. در سال‌های اخیر، روش‌های مبتنی بر یادگیری عمیق به دلیل دقت و انعطاف‌پذیری بالا، توجه بیشتری را به خود جلب کرده‌اند.

بحث و نتیجه‌گیری

ادغام تصویر به عنوان یک تکنیک حیاتی در پردازش تصویر، نقش بسیار مهمی در بهبود کیفیت و دقت داده‌های بصری ایفا می‌کند. با توجه به پیشرفت‌های سریع در فناوری و ظهور الگوریتم‌های جدید، پیش‌بینی می‌شود که این تکنیک در حوزه‌های مختلف علمی و صنعتی گسترش یابد و به کاربردهای نوینی دست یابد. درک عمیق‌تر از فرآیندهای ادغام تصویر و روش‌های مختلف آن می‌تواند به محققان و متخصصان کمک کند تا بهبودهای بیشتری در کیفیت داده‌ها و تجزیه و تحلیل‌های آن‌ها ایجاد کنند. علاوه بر این، توسعه مستمر الگوریتم‌های پیچیده‌تر، به ویژه الگوریتم‌های مبتنی بر یادگیری ماشین و یادگیری عمیق، پتانسیل‌های بیشتری را برای بهبود فرآیندهای ادغام تصویر فراهم می‌کند. این پیشرفت‌ها می‌تواند به بهبودهایی در زمینه‌هایی مانند تشخیص پزشکی، تصویربرداری ماهواره‌ای و بینایی



کامپیوتری منجر شود که در آن‌ها دقت و جزئیات تصاویر بسیار حیاتی است. همچنین با افزایش تقاضا برای داده‌های با کیفیت بالا در بخش‌های مختلف، توانایی ادغام تصاویر به‌طور کارآمد و دقیق، برای دستیابی به نتایج بهینه، حتی بیشتر از پیش اهمیت خواهد یافت. درک این روش‌ها همچنین راه‌های جدیدی برای توسعه راه‌حل‌های تخصصی در راستای چالش‌های خاص در حوزه‌های مختلف فراهم می‌آورد که منجر به نوآوری در پردازش و تفسیر داده‌های بصری خواهد شد.

- Cheng, C., et al. (2023). FusionBooster: A Unified Image Fusion Boosting Paradigm. arXiv preprint arXiv:2305.05970.
- Wang, X. (2022). Interpolation and sharpening for image upsampling. 2022 2nd International Conference on Computer Graphics, Image and Virtualization (ICCGIV). IEEE.
- Lee, S.-h., Park, J.S., and Cho, N.I. (2018). A multi-exposure image fusion based on the adaptive weights reflecting the relative pixel intensity and global gradient. 2018 25th IEEE International Conference on Image Processing (ICIP). IEEE.
- Khandelwal, S., et al. (2020). Accelerating local laplacian filters on FPGAs. 2020 30th International Conference on Field-Programmable Logic and Applications (FPL). IEEE.
- Gharbia, R., et al. (2014). Image fusion techniques in remote sensing. arXiv preprint arXiv:1403.5473.
- Karami, E., Prasad, S., and Shehata, M. (2017). Image matching using SIFT, SURF, BRIEF and ORB: Performance comparison for distorted images. arXiv preprint arXiv:1710.02726.
- Liu, J., et al. (2021). Learning a deep multi-scale feature ensemble and an edge-attention guidance for image fusion. IEEE Transactions on Circuits and Systems for Video Technology. 32(1), 105-119.
- Tasci, E., Uluturk, C., and Ugur, A. (2021). A voting-based ensemble deep learning method focusing on image augmentation and preprocessing variations for tuberculosis detection. Neural Computing and Applications. 33(22), 15541-15555.
- Grätzel von Grätz, P. (2021). Mehr Therapiespielraum für Patienten mit hereditärem Angioödem. Pädiatrie. 33(4), 62-64.
- Zhang, Y., et al. (2020). IFCNN: A general image fusion framework based on convolutional neural network. Information Fusion. 54, 99-118.
- Xu, H., Ma, J., and Zhang, X.-P. (2020). MEF-GAN: Multi-exposure image fusion via generative adversarial networks. IEEE Transactions on Image Processing. 29, 7203-7216.
- Li, H., Wu, X.-j., and Durrani, T.S. (2019). Infrared and visible image fusion with ResNet and zero-phase component analysis. Infrared Physics & Technology. 102, 103039.
- Yang, Z., et al. (2021). GANFuse: A novel multi-exposure image fusion method based on generative adversarial networks. Neural Computing and Applications. 33, 6133-6145.



Review of Techniques and Recent Advances in Image Fusion

Fatemeh Jokar

Instructor, Department of Computer Engineering, Shahid Bahonar University, Shiraz

MohammadReza Bahrani

Instructor, Department of Computer Engineering, Shahid Bahonar University, Shiraz

Seyed Arash Mousavi

Bachelor student, Department of Computer Engineering, Amirkabir University, Tehran

Abstract

Image fusion is a process that combines data from multiple images of a scene to create a more comprehensive and higher-quality image. This technique has numerous applications in fields such as remote sensing, medical imaging, and computer vision. Initially, traditional methods like weighted averaging and Laplace filters were used for image fusion, but with advancements in technology, deep learning-based methods such as Convolutional Neural Networks (CNN) and Generative Adversarial Networks (GAN) have been employed, significantly improving the accuracy and quality of image fusion.

These advancements have addressed many challenges, such as preserving image details, reducing noise, and optimizing processing. In particular, new algorithms have contributed to reducing computational complexity and improving accuracy in image fusion. Moreover, these techniques are applied in a wide range of applications, including disease detection in medical imaging, enhancing the quality of satellite images in remote sensing, and improving object detection accuracy in computer vision. This paper explores various image fusion methods, the challenges, and recent advancements in the field.

Keywords: Image fusion, medical imaging, remote sensing, image merging, image quality.