

:Viola-Jones

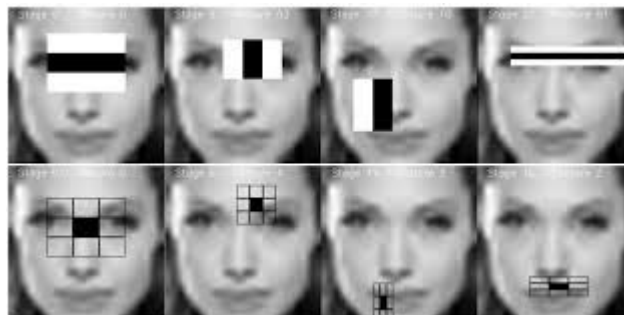
الگوریتم **Viola-Jones** یکی از نخستین و مهم‌ترین روش‌ها در حوزه بینایی ماشین برای تشخیص چهره است که در سال ۲۰۰۱ توسط "Paul Viola" و "Michael Jones" ارائه شد. این الگوریتم به دلیل سرعت بالا و دقت مناسبش، نقطه عطفی در توسعه سیستم‌های تشخیص چهره به حساب می‌آید و هنوز هم در بسیاری از پروژه‌ها و کتابخانه‌ها مثل OpenCV استفاده می‌شود. هدف اصلی این الگوریتم، تشخیص سریع اشیاء (به ویژه چهره‌ها) در تصاویر و ویدیوها است.

ایده اصلی این الگوریتم بر پایه استفاده از ویژگی‌های ساده تصویری به نام **ویژگی‌های Haar (Haar-like features)** است. این ویژگی‌ها تفاوت شدت روشنایی بین نواحی مختلف تصویر را اندازه‌گیری می‌کنند. برای مثال، اختلاف بین رنگ ناحیه چشم‌ها و گونه‌ها یا تفاوت میان بینی و اطرافش. این ویژگی‌ها بسیار ساده و محاسبه آن‌ها سریع است، اما به تنهایی نمی‌توانند چهره را دقیق تشخیص دهند؛ برای همین، از ترکیب تعداد زیادی از این ویژگی‌ها استفاده می‌شود.

برای اینکه محاسبه ویژگی‌های Haar سریع‌تر انجام شود، **Viola-Jones** از روشی به نام **تصویر انتگرالی (Integral Image)** استفاده می‌کند. این روش باعث می‌شود که محاسبه مجموع پیکسل‌های هر ناحیه از تصویر تنها با چند عملیات جمع ساده انجام شود. در نتیجه حتی اگر هزاران ویژگی Haar روی تصویر بررسی شود، زمان پردازش همچنان بسیار کم خواهد بود.

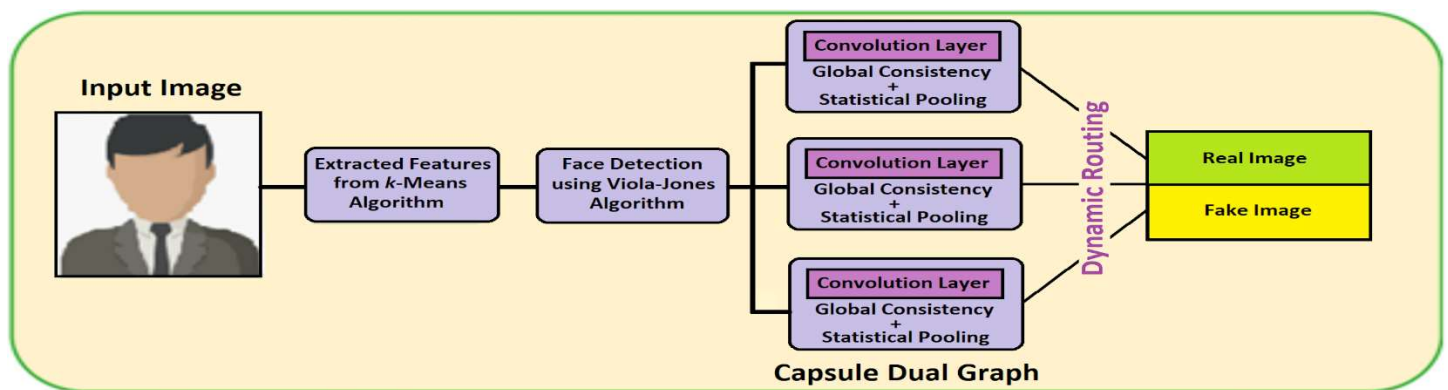
وقتی می‌خواهیم ویژگی‌های Haar را روی یک تصویر محاسبه کنیم، باید مجموع شدت پیکسل‌های یک ناحیه (مثلاً یک مربع یا مستطیل سفید و سیاه) را محاسبه کنیم. اگر تصویر بزرگ باشد و هزاران ویژگی Haar بررسی شود، محاسبه مجموع هر بار پیکسل‌ها بسیار زمان‌بر خواهد بود.

برای حل این مشکل، **Viola** و **Jones** از مفهوم **تصویر انتگرالی** استفاده کردند. در این روش، قبل از شروع محاسبه ویژگی‌ها، یک تصویر جدید ساخته می‌شود که در هر نقطه (x,y) مقدارش برابر است با مجموع همه پیکسل‌های بالاتر و سمت چپ آن نقطه. با داشتن تصویر انتگرالی، می‌توان مجموع شدت پیکسل‌های هر ناحیه مستطیلی را تنها با چهار عمل جمع و تفریق محاسبه کرد، فارغ از اندازه مستطیل!



این یعنی چه مستطیل 10×10 بخواهی حساب کنی، چه 100×100 ، تعداد عملیات همیشه ثابت است. ویژگی‌های Haar معمولاً به صورت اختلاف مجموع شدت پیکسل‌ها در دو یا چند ناحیه مجاور تعریف می‌شوند (مثل روشنایی چشم‌ها نسبت به گونه‌ها). اگر بخواهیم این ویژگی‌ها را بدون تصویر انتگرالی محاسبه کنیم، بسیار کند می‌شود. ولی با تصویر انتگرالی، محاسبه آن‌ها تقریباً آنی انجام می‌شود. پس **تصویر انتگرالی** همان چیزی است که الگوریتم Viola-Jones را از نظر سرعت بسیار کارآمد می‌کند. بدون آن، بررسی هزاران ویژگی Haar در هر فریم ویدئو امکان‌پذیر نبود. این ترنند باعث شد تشخیص چهره در زمان واقعی (Real-Time) روی سخت‌افزارهای قدیمی هم شدنی باشد.

از آنجایی که تعداد ویژگی‌های Haar بسیار زیاد است، باید انتخاب کنیم کدام یک واقعاً برای تشخیص مفید هستند. الگوریتم Viola-Jones برای این کار از تکنیک **AdaBoost** استفاده می‌کند. ویژگی‌های مؤثر را شناسایی کرده و ترکیب آن‌ها را به صورت یک "طبقه‌بند قوی" ارائه می‌دهد. به این ترتیب، ویژگی‌های بی‌اهمیت حذف می‌شوند و تنها ویژگی‌هایی باقی می‌مانند که بیشترین نقش را در تشخیص چهره دارند.



یکی از نقاط قوت الگوریتم Viola-Jones، استفاده از ساختار آبشاری (**Cascade**) است. به جای اینکه همه ویژگی‌ها را یکجا بررسی کند، الگوریتم آن‌ها را در چند مرحله پشت سر هم اعمال می‌کند. در هر مرحله، بخش زیادی از نواحی غیرچهره کنار گذاشته می‌شوند و تنها نواحی مشکوک برای بررسی‌های بعدی باقی می‌مانند. این ساختار باعث می‌شود تشخیص چهره بسیار سریع انجام شود، زیرا بیشتر قسمت‌های تصویر به سرعت حذف می‌شوند.

کتابخانه **OpenCV** پیاده‌سازی آماده‌ای از Viola-Jones را با نام **Haar Cascade Classifier** ارائه می‌دهد. در این روش، فایل‌های از پیش آموزش‌دیده (**XML**) وجود دارند که می‌توانند برای تشخیص چهره، چشم، لب‌خند یا حتی اشیاء دیگر استفاده شوند. این سیستم به راحتی قابل استفاده است و هنوز در بسیاری از پروژه‌های کاربردی (مثل سیستم‌های امنیتی یا اپلیکیشن‌های ساده موبایل) به کار می‌رود. هرچند روش‌های مدرن‌تر مثل شبکه‌های عصبی کانولوشنی (**CNN**) دقت بالاتری دارند، Haar Cascade همچنان به دلیل سبک بودن و سرعت بالا، کاربرد گسترده‌ای دارد.