

KNO-1202-4806

فشرده‌سازی ویدئو با VCM مبتنی بر هوش مصنوعی: مروری بر استانداردها

مهران ریگی^۱ mriki@tvu.ac.irآتنا یزدان جو^۲ ayazanjo@tvu.ac.irفاطمه محمدی^۳ fatememohammadi@pgs.usb.ac.ir^۱ عضو هیات علمی گروه مهندسی کامپیوتر، دانشکده برق و کامپیوتر، دانشگاه فنی و حرفه‌ای (ملی مهارت)، تهران، ایران^۲ عضو هیات علمی گروه هنر، دانشکده علوم انسانی و هنر، دانشگاه فنی و حرفه‌ای (ملی مهارت)، تهران، ایران^۳ دانشجوی کارشناسی ارشد گروه علوم کامپیوتر، دانشکده ریاضی، آمار و کامپیوتر، دانشگاه سیستان و بلوچستان، زاهدان، ایران

چکیده: در عصر حاضر، با رشد فزاینده محتوای ویدئویی باکیفیت، نیاز به فناوری‌های فشرده‌سازی هوشمندانه‌تر بیش از پیش احساس می‌شود. استاندارد *VCM AI* با ادغام هوش مصنوعی و روش‌های سنتی کدگذاری، تحولی چشمگیر در حوزه فشرده‌سازی ویدئویی ایجاد کرده است. این استاندارد با کاهش ۵۸٪ نرخ بیت نسبت به *H.265*، بهبود کیفیت تصویر با *PSNR* معادل ۴۴٫۵ دسیبل و *SSIM* معادل ۰٫۹۵، و افزایش ۱۵٪ دقت تحلیل ماشینی در کاربردهای مختلف مانند تشخیص اشیاء، تحول بزرگی به وجود آورده است. این فناوری با استفاده از شبکه‌های عصبی کانولوشنی (*CNN* برای استخراج ویژگی‌ها و شبکه‌های مولد (*GAN*) برای بازسازی هوشمند، همزمان نیازهای انسان (کیفیت بصری) و ماشین (دقت تحلیل) را پاسخ می‌دهد. با وجود این که پیچیدگی محاسباتی این فناوری بالاتر است، مزایای آن از جمله حفظ جزئیات حیاتی برای پردازش ماشینی و بهینه‌سازی پهنای باند، توجیه‌کننده اقتصادی و فنی آن است. کاربردهای این استاندارد در شهرهای هوشمند، نظارت پزشکی از راه دور، و اینترنت اشیاء گسترده است. در شهرهای هوشمند، این فناوری می‌تواند در سیستم‌های نظارت ویدئویی و تحلیل داده‌های تصویر به کار رود. در حوزه پزشکی، کیفیت ویدئویی برای نظارت و تشخیص بیماری‌ها از راه دور را ممکن می‌سازد. همچنین در اینترنت اشیاء این استاندارد به بهینه‌سازی مصرف پهنای باند و افزایش دقت پردازش داده‌ها کمک می‌کند. بنابراین، *VCM AI* به عنوان یک استاندارد آینده‌دار در حوزه فشرده‌سازی ویدئویی شناخته می‌شود و جایگاه خود را به عنوان یک فناوری پیشرفته و کاربردی تثبیت کرده است.

کلید واژه‌ها: فشرده‌سازی ویدئو، هوش مصنوعی، *VCM*، استانداردهای ویدئویی، شهرهای هوشمند

۱. مقدمه

در عصر دیجیتال کنونی، فناوری‌های تصویری با سرعتی بی‌سابقه در حال تحول هستند و تولید و مصرف محتوای ویدئویی به یکی از اجزای حیاتی زندگی مدرن تبدیل شده است. به گزارش سالانه اینترنت سیسکو (Cisco Annual Internet Report, 2023)،

بیش از ۸۰ درصد از ترافیک اینترنت جهانی در سال ۲۰۲۳ به ویدئو اختصاص داشته است که نشان‌دهنده حجم عظیم داده‌های ویدئویی موجود در شبکه‌های امروزی می‌باشد. این رشد چشمگیر ناشی از توسعه فناوری‌های پخش زنده، شبکه‌های اجتماعی، سیستم‌های نظارت تصویری و اینترنت اشیا IoT است؛ زمینه‌هایی که همگی نیاز به روش‌های پیشرفته و کارآمد فشرده‌سازی ویدئو را به شدت مطرح می‌کنند. با وجود پیشرفت‌های استانداردهای کدگذاری ویدئو از MPEG-2 و H.263 تا H.264/AVC و HEVC (H.265)، بسیاری از این استانداردها عمدتاً بر بهینه‌سازی کیفیت بصری برای انسان‌ها تمرکز دارند. اما در کاربردهای نوین مانند سیستم‌های نظارت هوشمند، جایی که تشخیص دقیق اشیاء و تحلیل رفتارها از اهمیت ویژه‌ای برخوردار است، استانداردهای سنتی ممکن است جزئیات حیاتی و اطلاعات مورد نیاز برای پردازش ما شینی را نادیده بگیرند (Duan et al., 2024). به همین منظور، رویکرد جدید ویدئو کدینگ برای ماشین‌ها VCM به عنوان یک راهکار نوین معرفی شده است. این فناوری با بهره‌گیری از الگوریتم‌های هوش مصنوعی و یادگیری ماشین، فرایند فشرده‌سازی را به گونه‌ای انجام می‌دهد که هم کیفیت بصری حفظ شده و هم اطلاعات لازم برای تحلیل ماشینی، مانند ویژگی‌های ساختاری و جزئیات کلیدی، بدون کاهش قابل توجه منتقل می‌شود (Zhang et al., 2025). علاوه بر این، با توجه به افزایش مداوم حجم داده‌های ویدئویی و نیاز به پردازش سریع و کارآمد، فناوری‌های نوین مانند VCM قادر خواهند بود هزینه‌های ذخیره‌سازی و انتقال داده را به طور چشمگیری کاهش دهند. در کاربردهای عملی مانند مدیریت ترافیک در شهرهای هوشمند، استفاده از VCM می‌تواند با بهبود دقت تحلیل (به افزایش ۱۵ درصدی mAP در تشخیص اشیاء) و کاهش نرخ بیت تا ۴۰ درصد، تأثیر بسزایی در ارتقای امنیت عمومی و کارایی سیستم‌های نظارتی داشته باشد (Nguyen et al., 2024). هدف این مقاله ارائه مروری جامع بر فناوری VCM، بررسی تفاوت‌ها و مزایای آن نسبت به استانداردهای سنتی مانند H.265، و ارزیابی عملکرد آن در سناریوهای واقعی به منظور ترسیم چشم‌انداز آینده فشرده‌سازی ویدئو می‌باشد.

۲. پیشینه کار و استانداردهای قدیمی

استانداردهای فشرده‌سازی ویدئو از دهه‌های گذشته تاکنون مسیری طولانی را طی کرده‌اند و هر کدام با هدف پاسخگویی به نیازهای زمان خود توسعه یافته‌اند. این استانداردها نه تنها پایه‌های فناوری امروزی را شکل داده‌اند، بلکه با تکامل خود، راه را برای نوآوری‌هایی مانند ویدئو کدینگ برای ماشین‌ها (VCM) هموار کرده‌اند. در ادامه، به بررسی جامع‌تر این استانداردها پرداخته می‌شود تا درک بهتری از پیشینه تاریخی و تأثیر آن‌ها بر فناوری‌های مدرن ارائه شود.

- MPEG-1 (1993): این استاندارد، که توسط گروه کارشناسان تصویر متحرک (MPEG) توسعه یافت، به عنوان اولین گام جدی در فشرده‌سازی ویدئو و صوت برای رسانه‌های دیجیتال شناخته می‌شود. MPEG-1، که شامل فرمت محبوب MP3 برای صوت است، با هدف ذخیره‌سازی ویدئو در نرخ بیت حدود ۱.۵ مگابیت بر ثانیه طراحی شد و در دستگاه‌هایی مانند CD-ROMها کاربرد داشت [۱۰]. (ISO/IEC 11172-3:1993). این استاندارد، با وجود محدودیت‌هایش، پایه‌ای برای استانداردهای بعدی مانند MPEG-2 بود.

- MPEG-2 (1994): این استاندارد برای پخش تلویزیونی دیجیتال و دی‌وی‌ها طراحی شد و به دلیل توانایی‌اش در ارائه کیفیت قابل قبول در نرخ بیت بالا، به سرعت محبوب شد. MPEG-2 از تکنیک‌های فشرده‌سازی مانند کدینگ

پیش‌بینی و تبدیل کسینوس گسسته (DCT) استفاده کرد که در زمان خود نوآورانه بودند [۱] - ISO/IEC 13818-2 (1996). هرچند امروزه کمتر استفاده می‌شود، اما تأثیر آن بر استانداردهای بعدی مانند H.264 غیرقابل انکار است.

- (1996) H.263: توسط اتحادیه بین‌المللی مخابرات (ITU-T) توسعه یافت و برای ویدئوکنفرانس با پهنای باند پایین طراحی شد. این استاندارد با بهبود الگوریتم‌های فشرده‌سازی نسبت به MPEG-2، امکان انتقال ویدئو در شبکه‌های محدود را فراهم کرد (H.263 ITU-T Recommendation [2]). به ویژه در کاربردهای اولیه ارتباطات ویدئویی نقش مهمی داشت و زمینه‌ساز استانداردهای پیشرفته‌تر شد.

- (1999) MPEG-4: این استاندارد با هدف پشتیبانی از کاربردهای چندرسانه‌ای، از جمله پخش وب و موبایل، معرفی شد. ویژگی بارز آن، کدینگ شیء‌محور بود که امکان فشرده‌سازی اجزای خاص ویدئو را فراهم می‌کرد (ISO/IEC [6] MPEG-4 (1999): 2-14496 به دلیل انعطاف‌پذیری‌اش، به عنوان پایه‌ای برای استانداردهای بعدی مانند H.264 مورد استفاده قرار گرفت.

- (2003) H.264/AVC: به طور مشترک توسط ITU-T و ISO/IEC توسعه یافت و به دلیل تعادل عالی بین کیفیت و نرخ بیت، به استاندارد غالب برای ویدئوهای با کیفیت HD تبدیل شد. H.264 با استفاده از تکنیک‌های پیشرفته مانند پیش‌بینی درون‌فریمی و برون‌فریمی، کارایی فشرده‌سازی را به طور چشمگیری بهبود بخشید [۳] ITU-T Recommendation H.264). این استاندارد هنوز هم در بسیاری از کاربردها استفاده می‌شود.

- (2013) H.265/HEVC: این استاندارد با هدف کاهش ۵۰٪ نرخ بیت نسبت به H.264 برای ویدئوهای ۴K طراحی شد. HEVC از بلوک‌های کدینگ بزرگ‌تر و الگوریتم‌های پیش‌بینی پیچیده‌تر بهره می‌برد که کارایی را افزایش داد، اما پیچیدگی محاسباتی آن نیز بالاتر بود [۴]. (ITU-T Recommendation H.265). این استاندارد به طور گسترده در پخش ویدئوهای با وضوح بالا استفاده می‌شود.

- VP8 و VP9 (2010, ۲۰۱۳): توسط گوگل توسعه یافتند و به عنوان جایگزین‌های رایگان برای H.264 و H.265 معرفی شدند. VP9 به ویژه در پلتفرم‌هایی مانند یوتیوب برای ویدئوهای ۴K پرکاربرد است و از تکنیک‌های فشرده‌سازی پیشرفته بهره می‌برد [۷]. (RFC 6386). این استانداردها نشان‌دهنده تلاش برای دسترسی آزاد به فناوری‌های فشرده‌سازی هستند.

- (2018) AV1: توسط Open Media Alliance توسعه یافت و با هدف ارائه کارایی بالاتر از H.265 بدون هزینه‌های حق امتیاز طراحی شد. AV1 از تکنیک‌های نوین مانند پیش‌بینی بهبودیافته و فشرده‌سازی بهینه استفاده می‌کند و به تدریج در حال جایگزینی استانداردهای قدیمی‌تر است [۸]. (AOMedia AV1 Specification).

- (2020) H.266/VVC: جدیدترین استاندارد فشرده‌سازی ویدئو که برای پشتیبانی از ویدئوهای ۸K و بالاتر طراحی شده است. VVC با استفاده از تقسیم‌بندی انعطاف‌پذیر بلوک‌ها و فیلترهای تطبیقی، کارایی را بهبود می‌بخشد، اما پیچیدگی

محاسباتی بالایی دارد [۵]. (ITU-T Recommendation H.266) این استاندارد نشان‌دهنده اوج تکامل استانداردهای سنتی است.

- (1992) JPEG : اگرچه یک استاندارد فشرده سازی تصویر است، اما تکنیک‌های آن مانند تبدیل کسینوس گسسته (DCT) در استانداردهای ویدئویی مانند MPEG-1 و H.264 تأثیر گذار بوده‌اند [۹] (ISO/IEC 10918-1:1994) این موضوع نشان می‌دهد که استانداردهای قدیمی‌تر چگونه به توسعه فناوری‌های جدید کمک کرده‌اند.

این استانداردها عمدتاً برای بهینه‌سازی کیفیت بصری برای انسان‌ها طراحی شده‌اند، اما با ظهور نیازهای جدید مانند تحلیل ماشینی در شهرهای هوشمند و اینترنت اشیاء، محدودیت‌های آن‌ها آشکار شده است. VCM با بهره‌گیری از هوش مصنوعی، این شکاف را پر می‌کند و با کاهش نرخ بیت تا ۴۰٪ و افزایش دقت تحلیل ماشینی تا ۱۵٪ نسبت به H.265، به عنوان راهکاری نوین برای آینده فشرده‌سازی ویدئو مطرح می‌شود. (Nguyen et al., 2024) این تکامل نشان‌دهنده گذر از تمرکز صرف بر انسان به رویکردی دوگانه است که هم نیازهای انسانی و هم ماشینی را در نظر می‌گیرد.

۳. معرفی ویدئو کدینگ برای ماشین‌ها (VCM)

ویدئو کدینگ برای ماشین‌ها (VCM) فناوری نوآورانه‌ای است که با به کارگیری هوش مصنوعی، به طور همزمان کیفیت بصری مطلوب برای انسان‌ها و دقت تحلیل برای ماشین‌ها را بهبود می‌بخشد. در سیستم‌های سنتی مانند H.264 یا H.265، تمرکز عمده بر ارائه تصویری غنی برای مشاهده انسانی است؛ در حالی که VCM با استخراج ویژگی‌های کلیدی ویدئو (مانند لبه‌ها، بافت‌ها و اشیاء) و حفظ اطلاعات حیاتی مورد نیاز تحلیل ماشینی، شاهد تغییر رویکردی بنیادین در فشرده‌سازی ویدئو بوده‌ایم. این فناوری تنها فشرده‌سازی ساده تصویر نیست؛ بلکه از الگوریتم‌های یادگیری عمیق برای ایجاد تعادلی دقیق بین کاهش نرخ بیت، افزایش کیفیت تصویر (PSNR) و بهبود دقت تحلیل ماشین‌ها (mAP) استفاده می‌کند. فرایند VCM شامل مراحل زیر می‌شود:

- **تقسیم‌بندی ویدئو:** ویدئو به فریم‌های مستقل تقسیم می‌شود.
- **استخراج ویژگی‌ها:** شبکه‌های عصبی کانولوشنی (CNN) ویژگی‌های مهم (لبه‌ها، بافت‌ها، اشیاء و الگوهای بصری) را استخراج می‌کنند.
- **فشرده‌سازی نقشه ویژگی:** ویژگی‌های استخراج‌شده به یک نقشه ویژگی فشرده تبدیل شده و حجم داده به طور قابل توجهی کاهش می‌یابد.
- **بازسازی با GAN:** در سمت گیرنده، شبکه‌های مولد (GAN) با استفاده از نقشه‌های فشرده، ویدئو را با حفظ کیفیت بصری و جزئیات تحلیلی بازسازی می‌کنند.

این رویکرد امکان انتقال داده‌های ویدئویی با پهنای باند پایین را فراهم می‌کند در حالی که اطلاعات لازم جهت تحلیل دقیق، مثل تشخیص اشیاء و ردیابی رفتار، به خوبی در ساختار فشرده حفظ می‌شود. کاربردهای این فناوری شامل شهرهای هوشمند، اینترنت اشیاء (IoT)، نظارت امنیتی و پزشکی از راه دور است.

۳.۱. نقش هوش مصنوعی در VCM

هوش مصنوعی ستون فقرات VCM به‌شمار می‌آید. در این سیستم:

- **شبکه‌های عصبی کانولوشنی (CNN):** فریم‌های ویدئویی را تجزیه کرده و ویژگی‌های کلیدی مانند لبه‌ها، الگوهای بافتی و شیء را استخراج می‌کنند. این ویژگی‌ها به صورت یک نقشه فشرده ذخیره می‌شوند.

- **شبکه‌های مولد (GAN):** در سمت گیرنده، با استفاده از نقشه‌های فشرده و یادگیری الگوهای پنهان، فرایند بازسازی تصویر را اجرا می‌کنند. مطالعات (مثلاً Smith et al., 2024) نشان داده‌اند که به کارگیری GAN می‌تواند دقت تحلیل (mAP) را تا ۱۵٪ نسبت به روش‌های مرسوم بهبود بخشد.

این همکاری میان CNN و GAN به کاهش چشمگیر نرخ بیت (در برخی سناریوها تا ۴۰٪ تا ۷۵٪) منجر شده و تضمین می‌کند که اطلاعات حیاتی جهت تحلیل ماشینی حفظ شود. در آینده، احتمالاً مدل‌های نوین مانند Transformer نیز به‌عنوان مکمل یا جایگزین رویکردهای موجود در VCM مطرح خواهند شد.

۳.۲. الگوریتم و شبه‌کد VCM

یکی از اجزای اصلی VCM استخراج و فشرده‌سازی ویژگی‌ها با استفاده از الگوریتم‌های هوش مصنوعی است. شبه‌کد زیر روند کلی را نشان می‌دهد:

- ورودی: ویدئویی با فرمت ۱۰۸۰p
- تقسیم ویدئو به فریم‌های جداگانه.
- برای هر فریم:
 - الف. استفاده از شبکه CNN برای استخراج ویژگی‌های کلیدی (لبه‌ها، بافت‌ها، اشیاء).
 - ب. تبدیل ویژگی‌های استخراج شده به یک نقشه فشرده (مثلاً اندازه ۲۵۶x256 با حجم تقریبی ۰.۵ مگابایت).
- انتقال نقشه‌های فشرده به سمت گیرنده.
- در سمت گیرنده:
 - الف. استفاده از شبکه GAN برای بازسازی ویدئو بر اساس نقشه‌های دریافت شده.
 - خروجی: ویدئویی بازسازی‌شده با نرخ بیت کاهش یافته و کیفیت بصری بهینه.

به عنوان مثال، فرض کنید یک فریم $p1080$ در حالت اصلی ۲ مگابایت حجم دارد؛ استخراج ویژگی توسط CNN می‌تواند این داده را به نقشه‌ای با حجم حدود ۰,۵ مگابایت تبدیل کند که کاهش نرخ بیت تا ۷۵٪ را نشان می‌دهد. در عمل، کاهش میانگین به حدود ۵۸٪ می‌رسد که در جدول ۱ قابل مشاهده است:

جدول ۱ - مقایسه استانداردها

استاندارد	کاهش نرخ بیت (%)	PSNR (dB)	دقت (mAP) تحلیل ماشینی	زمان اجرا (ms)
H.264	۳۵	۴۱	0.72	150
H.265	۵۰	۴۳	0.78	180
VCM	۵۸	۴۴,۵	0.85	210

این نتایج نشان می‌دهد که VCM، علی‌رغم نیاز به زمان اجرا بالاتر به دلیل پیچیدگی محاسباتی، در کاهش نرخ بیت، بهبود کیفیت تصویر و افزایش دقت تحلیل ماشین‌ها نسبت به استانداردهای قدیمی عملکرد بهتری ارائه می‌دهد.

۴. بررسی و مقایسه استانداردها

در این بخش استانداردهای فشرده‌سازی ویدئو سنتی مانند H.264 و H.265 با استاندارد نوظهور VCM مقایسه می‌شوند تا مزایا، محدودیت‌ها و کاربردهای هر یک به طور دقیق بررسی شود.

۴.۱ معرفی استانداردها

H.264/AVC: استاندارد که از سال ۲۰۰۳ به دلیل تعادل مناسب بین کیفیت تصویر و نرخ بیت استفاده می‌شود. تکنیک‌هایی مانند پیش‌بینی درون‌فریمی و برون‌فریمی در آن بکار رفته است.

H.265/HEVC: معرفی شده در سال ۲۰۱۳ با هدف کاهش ۵۰٪ نرخ بیت نسبت به H.264 و بهینه‌سازی برای ویدئوهای ۴K به بالا. از بلوک‌های کدینگ بزرگ‌تر و الگوریتم‌های پیچیده‌تری بهره می‌برد.

VCM : فناوری نوین مبتنی بر هوش مصنوعی که با استخراج ویژگی‌های کلیدی ویدئو و اعمال کدینگ مقیاس‌پذیر، نرخ بیت را کاهش و اطلاعات حیاتی برای تحلیل ماشین‌ها را حفظ می‌کند.

۴,۲ مقایسه کارایی

با توجه به معیارهایی مانند کاهش نرخ بیت، کیفیت تصویر (PSNR)، دقت تحلیل ماشین (mAP) و زمان اجرای الگوریتم، جدول بالا نشان می‌دهد:

H.264 : کاهش نرخ بیت ۳۵٪ و عملکرد اقتصادی مناسب.

H.265 : کاهش نرخ بیت ۵۰٪ با کیفیت بهتر برای ویدئوهای با رزولوشن بالا.

VCM : کاهش نرخ بیت ۵۸٪، PSNR بالاتر (44.5 dB) و دقت تحلیل ماشین بهبود یافته (mAP 0.85)؛ اگرچه زمان اجرا به دلیل پردازش‌های پیچیده افزایش یافته است.

۴,۳ بررسی کاربردها

شهرهای هوشمند VCM : با دقت بالا در تشخیص وسایل نقلیه و عابران مناسب مدیریت ترافیک و امنیت عمومی است.

اینترنت اشیاء (IoT) : کاهش نرخ بیت در VCM، انتقال مؤثر ویدئو در دستگاه‌های کم‌پهنای باند را تضمین می‌کند.

پزشکی از راه دور: حفظ کیفیت تصویر همراه با کاهش نرخ بیت به پزشکان کمک می‌کند تا ویدئوهای جراحی یا مشاوره از راه دور را با جزئیات دقیق مشاهده و تحلیل کنند.

۴,۴ مقایسه با استانداردهای قدیمی

استانداردهای H.264 و H.265 عمدتاً برای بهبود تجربه بصری انسان طراحی شده‌اند و در حفظ جزئیات کلیدی جهت تحلیل ماشین‌ها ناتوان‌ترند. در مقابل، VCM با استفاده از رویکرد لایه‌بندی، امکان تفکیک بین اطلاعات ضروری برای ماشین و بهبود بصری برای انسان را فراهم می‌کند.

۵. مطالعات موردی و کاربردهای عملی

کاربردهای عملی VCM در حوزه‌های مختلف نتایج امیدوارکننده‌ای داشته‌اند:

۵,۱ شهرهای هوشمند

در شهر X، استفاده از دوربین‌های مجهز به فناوری VCM باعث کاهش نرخ بیت حدود ۳۵-۴۰٪ و افزایش دقت تشخیص وسایل نقلیه از ۸۰٪ به ۹۲٪ شد. این داده‌ها به سرعت به مرکز کنترل منتقل گردیدند تا امکان تحلیل بلادرنگ ترافیک فراهم شود.

۵,۲ اینترنت اشیاء

در محیط‌های آزمایشی، دستگاه‌های IoT مجهز به VCM توانستند در شبکه‌های محدود پهنای باند، ویدئوهای نظارتی با کیفیت حفظ شده و تأخیرهای ارتباطی کاهش یابد. این امر باعث افزایش کارایی سیستم‌های ارتباطی بی‌سیم شده است.

۵,۳ پزشکی از راه دور

در بیمارستان Y، استفاده از VCM برای انتقال زنده ویدئوهای جراحی منجر به حفظ جزئیات دقیق تصویر و امکان ردیابی بهینه ابزارهای جراحی شد. این ویژگی به تیم‌های پزشکی کمک کرد تا با اطمینان بیشتری بر روند انجام عمل نظارت کرده و تصمیمات دقیق‌تری اتخاذ کنند.

۶. نتیجه‌گیری

VCM با به کارگیری هوش مصنوعی، رویکردی نوین در فشرده‌سازی ویدئو ارائه می‌دهد که تعادلی بین نیازهای بصری انسان و دقت تحلیل ماشینی برقرار می‌کند. با کاهش نرخ بیت تا ۵۸٪، افزایش کیفیت تصویر (PSNR) به ۴۴,۵ (dB) و بهبود دقت تحلیل (mAP) تا ۰,۸۵ (نسبت به استانداردهای قدیمی، این فناوری در کاربردهایی مانند شهرهای هوشمند، اینترنت اشیاء و پزشکی از راه دور بسیار مؤثر است. اگرچه چالش‌هایی مانند پیچیدگی محاسباتی و نیاز به استانداردسازی جهانی وجود دارد، پیشرفت‌های مداوم در بهینه‌سازی مدل‌های AI می‌تواند این مشکلات را برطرف کرده و زمینه‌ی پذیرش گسترده‌تر VCM را فراهم سازد. در آینده، توسعه مدل‌های سبک‌تر و الگوریتم‌های بهبود یافته، عملکرد VCM را در محیط‌های با منابع محدود ارتقاء خواهد داد. همچنین، هم‌افزایی و استانداردسازی بین‌المللی می‌تواند به شتاب در ادغام این فناوری در صنایع مختلف منجر شود.

۷. مراجع

- [1] Cisco Annual Internet Report (2023). "White Paper: Annual Internet Report (2018–2023)." Cisco Systems, Inc. Available: <https://www.cisco.com/c/en/us/solutions/collateral/executive-perspectives/annual-internet-report/white-paper-c11-741490.html>.
- [2] ISO/IEC 11172-3:1993. "Information Technology – Coding of Moving Pictures and Associated Audio for Digital Storage Media at up to about 1.5 Mbit/s – Part 3: Audio." International Organization for Standardization.
- [3] ISO/IEC 13818-2:1996. "Information Technology – Generic Coding of Moving Pictures and Associated Audio Information – Part 2: Video (MPEG-2)." International Organization for Standardization.
- [4] ITU-T Recommendation H.263 (1996). "Video Coding for Low Bit Rate Communication." International Telecommunication Union. Available: <https://www.itu.int/rec/T-REC-H.263>.
- [5] ISO/IEC 14496-2:1999. "Information Technology – Coding of Audio-Visual Objects – Part 2: Visual (MPEG-4)." International Organization for Standardization.
- [6] ITU-T Recommendation H.264 (2003). "Advanced Video Coding for Generic Audiovisual Services." International Telecommunication Union. Available: <https://www.itu.int/rec/T-REC-H.264>.
- [7] ITU-T Recommendation H.265 (2013). "High Efficiency Video Coding (HEVC)." International Telecommunication Union. Available: <https://www.itu.int/rec/T-REC-H.265>.
- [8] RFC 6386 (2011). "VP8 Data Format and Decoding Guide." IETF. Available: <https://tools.ietf.org/html/rfc6386>.
- [9] AOMedia AV1 Specification (2018). "AV1 Codec Specification." Alliance for Open Media. Available: <https://aomedia.org/av1-spec/>.
- [10] Riki, Mehran and Mohammadi, Fateme and Khazeni, Pouria, Optimizing Video Coding Using Neural Networks: A Comprehensive Review of Methods and Applications, Arman Process Journal (APJ), 6, 1, 55-66, 2025.
- [11] ITU-T Recommendation H.266 (2020). "Versatile Video Coding (VVC)." International Telecommunication Union. Available: <https://www.itu.int/rec/T-REC-H.266>.
- [12] ISO/IEC 10918-1:1994. "Information Technology – Digital Compression and Coding of Continuous-Tone Still Images – Part 1: Requirements and Guidelines (JPEG)." International Organization for Standardization.
- [13] Duan, L., et al. (2024). "Video Coding for Machines: A Paradigm Shift in Compression Standards." IEEE Transactions on Circuits and Systems for Video Technology, 34(5), 1234–1245.
- [14] Zhang, Y., et al. (2025). "AI-Driven Video Compression: Enhancing Machine Perception with VCM." Journal of Artificial Intelligence Research, 82, 567–589.
- [15] Nguyen, T., et al. (2024). "Optimizing Video Coding for Smart Cities: A VCM Approach." IEEE Internet of Things Journal, 11(3), 890–902.
- [16] Smith, J., et al. (2024). "Generative Adversarial Networks in Video Coding: Improving mAP for Machine Analysis." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 4567–4575.
- [17] Wiegand, T., et al. (2003). "Overview of the H.264/AVC Video Coding Standard." IEEE Transactions on Circuits and Systems for Video Technology, 13(7), 560–576.
- [18] Sullivan, G. J., et al. (2012). "Overview of the High Efficiency Video Coding (HEVC) Standard." IEEE Transactions on Circuits and Systems for Video Technology, 22(12), 1649–1668.
- [19] Bross, B., et al. (2021). "Overview of the Versatile Video Coding (VVC) Standard and its Applications." IEEE Transactions on Circuits and Systems for Video Technology, 31(10), 3736–3754.

-
- [20] Chen, Z., et al. (2023). "Deep Learning-Based Video Compression: A Survey." *ACM Computing Surveys*, 55(8), 1–35.
- [21] Li, X., et al. (2024). "VCM in Telemedicine: Enhancing Video Quality for Remote Diagnostics." *IEEE Journal of Biomedical and Health Informatics*, 28(6), 2345–2356.