ژو و همکاران

\cite{CoOp}،

رویکردی به نام به بهینه‌سازی بافت

\LTRFootnote{Context Optimization (CoOp)}

ارائه کرده‌اند که به جای پرامپت‌های ثابت، از پرامپت به صورت بردارهای بافت یادگرفتنی

\LTRFootnote{Learnable context vectors}

که در کنار برچسب متنی داده‌ها قرار می‌گیرند، استفاده می‌کند. مطابق

\cref{fig2.3}،

تمام وزن‌های مدل

\lr{CLIP}،

ثابت نگه داشته شده و تنها بردارهای پرامپت، قابل‌آموزش هستند. به دنبال توسعه ی روش های پرامپت گذاری، روش

\lr{L2P}

\LTRfootnote{Learning to Prompt for Continual Learning}

که توسط وانگ و همکاران

\cite{l2p}

ارائه شده است، یک چارچوب نوآورانه برای یادگیری پیوسته بدون نیاز به شناسایی وظیفه در زمان آزمون می‌باشد. همان‌طور که در

\cref{fig.23}

مشاهده می‌شود، این روش به جای تغییر وزن‌های مدل پیش‌آموخته، از مجموعه‌ای از پرامپت‌های یادگرفتنی بهره می‌برد که در یک فضای حافظه اشتراکی به نام استخر پرامپت

\LTRfootnote{Prompt pool}

نگهداری می‌شوند.

\lr{L2P}

از یک مکانیزم پرس‌وجوی مبتنی بر جفت‌های کلید-مقدار بهره می‌برد تا به‌صورت پویا و متناسب با ورودی، پرامپت‌های مرتبط را انتخاب کرده و به نشانه‌‌های ورودی مدل، اضافه کند. سپس این نشانه‌‌های توسعه‌یافته به مدل پیش‌آموخته تزریق شده و پیش‌بینی انجام می‌شود.

در این روش، پرامپت‌ها، دانش خاص هر وظیفه یا دانش مشترک بین وظایف را به‌صورت فشرده ذخیره می‌کنند و باعث کاهش چشمگیر فراموشی مخرب در یادگیری وظایف متوالی می‌شوند. ساختار طراحی‌شده در

\lr{L2P}

همان‌طور که در

\cref{fig.23}

نشان داده شده، از یک بخش انتخاب پرامپت، لایه‌های کدگذار پیش‌آموخته

\lr{Pretrained encoder layers}،

و دسته‌بند نهایی تشکیل شده است. جهت بهبود و مقاوم‌سازی نسبت به فراموشی در انتخاب پرامپت در روش‌های پیشین، مارتین و همکاران

\cite{starPrompt}،

روش

\lr{STAR-Prompt}

را معرفی کرده اند که از رویکردی دوسطحی برای تنظیم پرامپت، پیروی می‌کند. ابتدا از

\lr{CLIP}،

برای تولید پرامپت‌های متنی و ساخت نمونه‌های‌ اولیه

\LTRfootnote{Prototypes}،

پایدار دسته‌ها استفاده می‌شود و سپس این نمونه‌های اولیه به‌عنوان کلید برای بازیابی پرامپت‌های تصویری در ترنسفورمر تصویر به‌کار می‌روند. هم‌چنین روش‌های

\lr{DualPrompt}

\cite{dualprompt}

و

\lr{H-prompts}

\cite{h-prompts}،

مانند مطالعات مذکور، در زمینه‌ی تولید پرامپت‌های مشترک و خاص وظایف ارائه شده‌اند. داهویین و همکاران

\cite{instance\_prompt}،

از پرامپت اختصاصی برای هر نمونه به جای هر دسته، استفاده کرده‌اند. برخلاف روش های پرامپت گذاری قبلی، هنگ و همکاران

\cite{ovor}،

علاوه بر استفاده از یک پرامپت به‌جای چندپرامپت، از نمونه‌های پرت مصنوعی برای ایجاد مرز دسته‌بندی بهتر استفاده کرده‌اند. در مطالعات دیگری مانند روش یکپارچه‌سازی دانش بدون تداخل و آگاه از توزیع

\cite{diki}

\LTRfootnote{Distribution-aware Interference-free Knowledge Integration (DIKI)}،

به رفع چالش مداخله‌ی پرامپت در تصمیم گیری مکانیزم توجه پرداخته شده است. به دنبال پیشرفت های روش های پرامپت در حوزه‌ی تصویر، ویلا و همکاران

\cite{pivot}،

روشی به نام

\lr{PIVOT}

را معرفی کرده‌اند که با بهره‌گیری از دانش پیش‌آموخته مدل‌ تصویر-متن

\lr{CLIP}

و استفاده از پرامپت‌های مکانی

\LTRfootnote{Spatial prompts}

و زمانی، وابستگی‌های زمانی و مکانی ویدیوها را مدل کرده است.

.

به دنبال پیشرفت های روش های پرامپت در حوزه ی تصویر، pivot پرامپت زمانی برای سازگار کردن مدل های تصویری به ویدیو ارائه داد. مقاله ی seeing the flowing نیز پرامپت حرکتی و غیره. در vilt-clip نیز پرامپت چند ماهیتی ارائه شده.

@article{CoOp,

title={Learning to Prompt for Vision-Language Models},

author={Zhou, Kaiyang and Yang, Jingkang and Loy, Chen Change and Liu, Ziwei},

journal={International Journal of Computer Vision (IJCV)},

year={2022}

}