دوفصل‌نامه علمی منادی امنیت فضای تولید و تبادل اطلاعات (افتا)
سال ۱۴۰۰، جلد ۱۹، شماره ۱ (۴۳-۵۱)
https://monadi.isc.org.ir

ارائه‌شده در شانزدهمین کنفرانس بین المللی انجمن رمز ایران در دانشگاه فردوسی مشهد، ۶ و ۷ شهریور ۱۳۹۸

کاربرد یادگیری عمیق و شبکه عصبی پیچشی در نهان‌کاوی★

مهدیه سمیعی* و وجیهه ثابتی

گروه مهندسی کامپیوتر، دانشگاه الزهرا، تهران، ایران

اطلاعات مقاله

کلمات کلیدی:
نهان‌کاوی
نهان‌نگاری
شبکه‌های عصبی
یادگیری عمیق
شبکه‌های عصبی پیچشی

dor: 10.0000/000000000

نوع مقاله: پژوهشی

چکیده

نهان‌نگاری، ابزاری برای ارتباط محرمانه و در مقابل نهان‌کاوی علم کشف حضور اطلاعات نهان در رسانه دیجیتال می‌باشد. تاکنون نهان‌کاوی تصاویر دیجیتالی روی ویژگی‌های دست‌ساز پیچیده متمرکز بوده‌اند که از جمله آن می‌توان به مدل معروف و موفق SRM اشاره کرد، اما امروزه با استفاده از مدل‌های یادگیری عمیق می‌توان ویژگی‌های را به صورت خودکار استخراج کرد به عبارت دیگر مراحل استخراج ویژگی و طبقه‌بندی تحت یک معماری واحد قرار گرفتند. تکنیک‌های نهان‌کاوی مختلفی در تصاویر با استفاده از الگوریتم‌های یادگیری عمیق از جمله شبکه‌های عصبی پیچشی پیاده سازی شده‌اند. در این مقاله، به معرفی چهار تکنیک نهان‌کاوی GNCNN، Xu-NET، Ye-NET، YEDROUDJ-NET در این حوزه پرداخته شده است. پس از بررسی و مقایسه نتایج این چهار روش مشاهده شد که تکنیک YEDROUDJ-NET توانسته است به خطای احتمالی مشابه و در اغلب موارد کمتر از مدل SRM دست یابد. بنابراین روش‌های نهان‌کاوی مبتنی بر شبکه‌های عصبی پیچشی توانسته‌اند کارایی مشابه و حتی در مواردی بهتر از روش‌های نهان‌کاوی سنتی ارائه دهند.

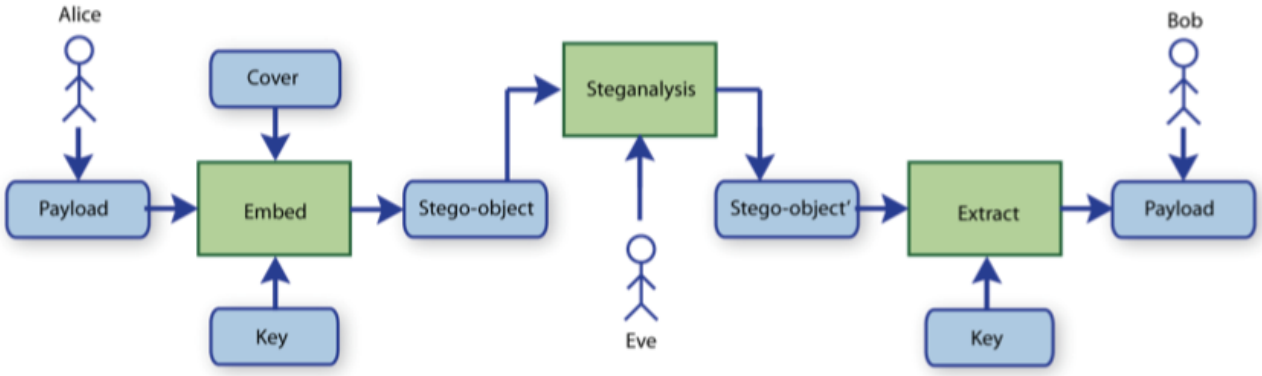
۱ مقدمه

با پیشرفت در فناوری ارتباطات دیجیتال و افزایش قدرت و ذخیره‌سازی کامپیوترها، حفظ حریم خصوصی افراد به یک چالش مهم تبدیل شده است. میزان حریم خصوصی از شخصی به شخص دیگر متفاوت است [۱]. نهان‌سازی داده‌ها مجموعه‌ای از تکنیک‌های جاسازی اطلاعات مخفیانه در رسانه‌های دیجیتال است. این تکنیک‌ها می‌توانند در بسیاری از سناریوهای کاربردی مختلف مانند ارتباطات محرمانه، حفاظت از حق تکثیر یا احراز هویت محتوای دیجیتال مورد استفاده قرار گیرند. از نهان‌نگاری به عنوان یکی از روش‌های نهان‌سازی داده‌ها برای کاربرد فراهم کردن ارتباط محرمانه استفاده می‌شود. در این روش تصاویر یکی از رایج‌ترین حامل‌های اطلاعات مخفیانه به دلیل استفاده گسترده از آن‌ها در اینترنت می‌باشند [۲]. در مقابل نهان‌کاوی [1] مهارت کشف داده‌های نهان و یا وجود داده‌های نهان است.

تکنیک‌های نهان‌کاوی نه تنها به روش‌های جاسازی اطلاعات، بلکه به نوع رسانه پوشانه‌ای [2] که توسط سیستم به کار گرفته می‌شود بستگی دارد. به طور کلی داده جاسازی شده در تصاویر طبیعی چالش سخت‌تری را برای تشخیص آن داده نسبت به تصاویر تولید شده توسط کامپیوتر نشان می‌دهد [۳]. این موضوع در درجه اول به دلیل افزایش واریانس در رنگ، تضاد، تکنیک و غیره می‌باشد و همچنین نویز ایجاد شده توسط دوربین‌ها تشخیص تصویر گنجانه [3] از مجموعه‌ای از تصاویر پوشانه را

[1] steganalysis [2] cover [3] stego

★از کمیته علمی شانزدهمین کنفرانس بین‌المللی انجمن رمز ایران برای داوری این مقاله تشکر می‌شود.
*نویسنده مسئول
آدرس‌های رایانامه: samiee.mahdis@yhoo.com (مهدیه سمیعی)، v.sabeti@alzahra.ac.ir (وجیهه ثابتی)



شکل ۱. فرایند کامل نهان‌نگاری، شامل نهان‌کاوی [۳]

سخت‌تر می‌کند. تکنیک‌های استفاده شده در نهان‌کاوی به طورگسترده‌ای متفاوت هستند، اما همه آن‌ها عموما در سه دسته آماری [۱]، بصری [۲] و یا ساختاری [۳] طبقه بندی می‌شوند که شرح آن‌ها به صورت زیر است [۳].

نهان‌کاوی بصری: ساده‌ترین شکل نهان‌نگاری، شناسایی ناهنجاری‌های بصری داخل تصویر گنجانه است. بسیاری از تکنیک‌های نهان‌نگاری بصری به نقص در الگوریتم‌های جاسازی متکی هستند. ساخت شیء گنجانه مشابه با اصل آن دیگر یک چالش نیست چون اجزای جدا شده می‌توانند نقایص را آشکار کنند. بنابراین تجزیه یک تصویر به عناصر تشکیل‌دهنده‌ی آن، اغلب برای تحلیل بصری روش‌های بسیار بهتری را نسبت به پردازش کل تصویر به صورت یکنواخت، به ارمغان می‌آورد.

نهان‌کاوی ساختاری: نهان‌کاوی ساختاری تغییرات ایجاد شده در فرمت فایل گنجانه را تشخیص می‌دهد و با مقایسه ساختار و استانداردی که برای آن تعریف شده حضور داده جاسازی شده نتیجه می‌شود. فرمت‌های تصویری که از پالت‌های رنگ استفاده می‌کنند برای تجزیه و تحلیل ساختاری آسیب پذیرتر هستند. به منظور مقابله با نقص‌های موجود در پیکسل‌های مربوط به ورودی اشتباه، تغییرات مربوط به ساختار پالت معرفی شده که با تجزیه و تحلیل ساختاری می‌توان آن‌ها را تشخیص داد.

نهان‌کاوی آماری: تجزیه و تحلیل ویژگی‌های آماری یک شیء گنجانه، در مقایسه با مجموعه‌ای از پوشانه‌ها، اغلب حضور داده‌های نهان را نشان می‌دهد. این تکنیک‌ها اغلب یک محیط جاسازی خاص را برای ایجاد یک پایه آماری توسط مجموعه‌ای از پوشانه‌ها، مورد هدف قرار می‌دهند. روند کامل نهان‌نگاری و نهان‌کاوی در شکل ۱ نشان داده شده است.

در حال حاضر امن‌ترین روش‌های نهان‌نگاری تمایل دارند تا داده را در قسمت‌های پیچیده تصویر نهان کنند. از جمله روش‌های نهان‌نگاری موفق در حوزه مکان عبارتند از: WOW [۴] و S-UNIWARD [۵].

با گسترده تر شدن روش‌های نهان‌نگاری پیشرفت قابل ملاحظه‌ای در تکنیک‌های نهان‌کاوی صورت گرفته که از جمله این تکنیک‌ها در حوزه مکانی می‌توان به مدل SRM اشاره کرد. اخیرا بهترین تکنیک‌های نهان‌کاوی بر مبنای ویژگی‌های به دست آمده از تصاویر و روش‌های یادگیری ماشین بوده که همه آن‌ها از یک الگوی مشترک استفاده می‌کنند: محاسبه نویز به دست آمده در تصویر، ساخت ویژگی‌ها و طبقه‌بند دودویی [۶]. در این روش‌ها، ابتدا ویژگی‌ها (ویژگی‌ها ممکن است در ابعاد بالا عرضه شوند) با استفاده از روش‌های مختلف، از مجموعه داده‌های آموزشی برای آموزش طبقه‌بند استخراج می‌شوند. سپس طبقه‌بند با استفاده از این مجموعه‌های آموزشی اعتبارسنجی می‌شود، اگر نتایج دقیق باشد، طبقه‌بند به عنوان یک طبقه‌بند مفید در نظر گرفته می‌شود [۷].

برخی از روش‌های استخراج ویژگی که اطلاعات مورد نیاز را از تصاویر اصلی استخراج می‌کنند عبارتند از: وابستگی خودکار نسبی (RAD [۴]) [۸]، تبدیل کسینوسی گسسته (DCT [۵]) [۹]، تبدیل موجک گابور [۶] [۱۰]، استخراج ویژگی های مبتنی بر طیف [۱۱]، کوانتیزاسیون نامنظم (PQ [۷]) [۱۲]، تبدیل کانتورلت بدون نمونه (NSCT [۸]) [۱۳]. چند نمونه از طبقه‌بندهایی که در نهان‌کاوی مورد استفاده قرار می‌گیرند عبارتند از: طبقه‌بند گروهی [۹] [۱۴]، طبقه‌بند EN_ELM [۱۰] [۱۵]، طبقه‌بند ماشین یادگیری سریع [۱۱] [۱۶]، طبقه‌بند دیفرانسیل تکاملی (DE [۱۲]) [۱۷]، طبقه‌بند گروهی شناختی ELM [۱۸]، طبقه‌بند ماشین بردار پشتیبان فازی [۱۳] [۱۹]، طبقه‌بند بیزین [۲۰] و غیره.

همانطور که گفته شد در این طبقه‌بندها ویژگی‌ها به صورت دستی به عنوان ورودی داده می‌شود، اما در این مقاله نشان داده شده که الگوی مشترک ذکر شده می‌تواند توسط شبکه عصبی پیچشی پیاده سازی شود. مهم‌ترین ویژگی شبکه عصبی پیچشی [۱۴] این است که می‌تواند وابستگی‌های آماری پیچیده را از ویژگی‌های با ابعاد بسیار بالا به صورت خودکار استخراج کند و به صورت عمیق برای دقیق‌تر شدن نتیجه طبقه‌بند آموزش ببیند.

سایر قسمت‌های مقاله به شرح زیر است: در بخش ۲ یک معرفی

[1] statistical [2] visual [3] structural

[4] Relative Auto Decorrelation [5] Discrete Cosine Transform [6] Gabor Wavelet Transform [7] Perturbed Quantization [8] Non-Sampled contourlet Transform [9] Ensemble classifier [10] Ensemble based Extreme Learning Machine [11] Extreme Learning Machine [12] Differential Evolution [13] fusing support vector machine [14] convolutional neural networks

اجمالی از یادگیری عمیق و شبکه‌های عصبی گفته شده است، در بخش ۳ یک نگاه کلی به معماری SRM ارائه شده است، در بخش ۴ به معرفی چهار مدل نهان‌کاوی که با استفاده از شبکه‌های عصبی پیچشی پیاده‌سازی شدند پرداخته شده، در بخش ۵ تحلیل و مقایسه نتایج این چهار تکنیک بررسی شده و در بخش ۶ به نتیجه‌گیری پرداخته شده است.

۲ شبکه‌های عصبی و یادگیری عمیق

شبکه‌های عصبی مصنوعی برگرفته شده از شبکه عصبی انسان هستند. از چند دهه گذشته که رایانه‌ها امکان پیاده‌سازی الگوریتم‌های محاسباتی را فراهم ساخته‌اند، در راستای شبیه‌سازی رفتار محاسباتی مغز انسان، کارهای پژوهشی بسیاری از سوی متخصصین علوم رایانه، مهندسین و همچنین ریاضی‌دان‌ها شروع شده است، که نتایج کار آن‌ها، در شاخه‌ای از علم هوش مصنوعی و در زیرشاخه هوش محاسباتی تحت عنوان موضوع شبکه‌های عصبی مصنوعی طبقه‌بندی شده است.

یادگیری عمیق یک نوع از یادگیری ماشین است که کامپیوترها را قادر می‌سازد تا از تجربه یاد بگیرند و مفاهیم جهان را به صورت سلسله مراتبی فراگیرند. سلسله مراتب به کامپیوترها کمک می‌کند تا مفاهیم پیچیده را با ساختن مفاهیم ساده‌تر از آن‌ها یاد بگیرند.

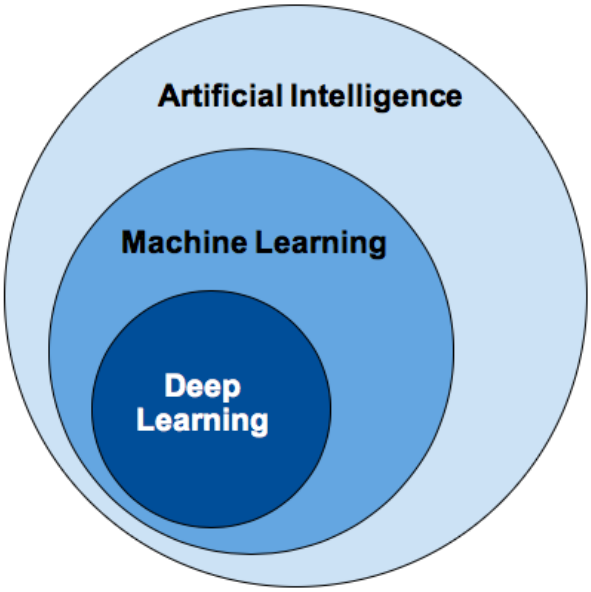
اگر یک گراف از این مؤلفه‌ها که به صورت سلسله مراتبی کنار همدیگر قرار گرفته‌اند، رسم شود، این گراف شامل چندین لایه عمیق خواهد بود. یادگیری عمیق در واقع همان شبکه عصبی عمیق (چندلایه) است که در آن از مدل‌های دیگری برای چینش نورون‌ها استفاده می‌شود که نحوه این چینش می‌تواند به صورت تمام متصل باشد که در آن تعداد لایه‌های نهان بیشتر است و هرچه بیشتر پیش می‌رویم ویژگی‌های بیش‌تر و پیچیده‌تری استخراج می‌شوند. همانطور که در شکل ۲ نشان داده شده است، یادگیری عمیق زیر مجموعه‌ای از یادگیری ماشین و یادگیری ماشین نیز زیر مجموعه‌ای از هوش مصنوعی است.

۳ نگاهی به SRM

در حال حاضر SRM [1] یکی از موفق‌ترین و جدیدترین مدل‌های مبتنی بر ویژگی است که بدون آگاهی از روش نهان‌نگاری تصاویر، گنجانه یا پوشانه بودن آن‌ها را تشخیص می‌دهد. این مدل از لحاظ ساختاری شباهت‌های زیادی به شبکه‌های عصبی پیچشی دارد، به همین دلیل تکنیک‌های مختلف نهان‌کاوی در یادگیری عمیق با مدل SRM مقایسه شده تا به نتایجی مشابه و یا بهتر دست یابند. این تکنیک‌ها در لایه‌های اول خود از کرنل‌های تولید شده از SRM استفاده می‌کنند.

ایده اصلی SRM این است که انواع مختلفی از وابستگی‌ها را در میان پیکسل‌های همسایه ثبت می‌کند، که این حالت تضمین سازگاری بالایی برای مدل به دست آمده را می‌دهد. وابستگی‌ها به عنوان مانع

[1] Spatial domain Rich Model



شکل ۲. ارتباط بین هوش مصنوعی، یادگیری ماشین و یادگیری عمیق

نویز مدل‌سازی می‌شوند. این نویز باقی‌مانده طبق فرمول (۱) محاسبه می‌شود که در آن X_{ij} پیکسلی در عکس X در محل (i,j)، N_{ij} همسایه محلی پیکسل X_{ij}، $\theta(N_{ij})$ پیش‌بینی‌کننده‌ی X_{ij} روی N_{ij}، و $R=(R_{ij}) \in R^{n_1 \times n_2}$ نویز باقی‌مانده می‌باشد:

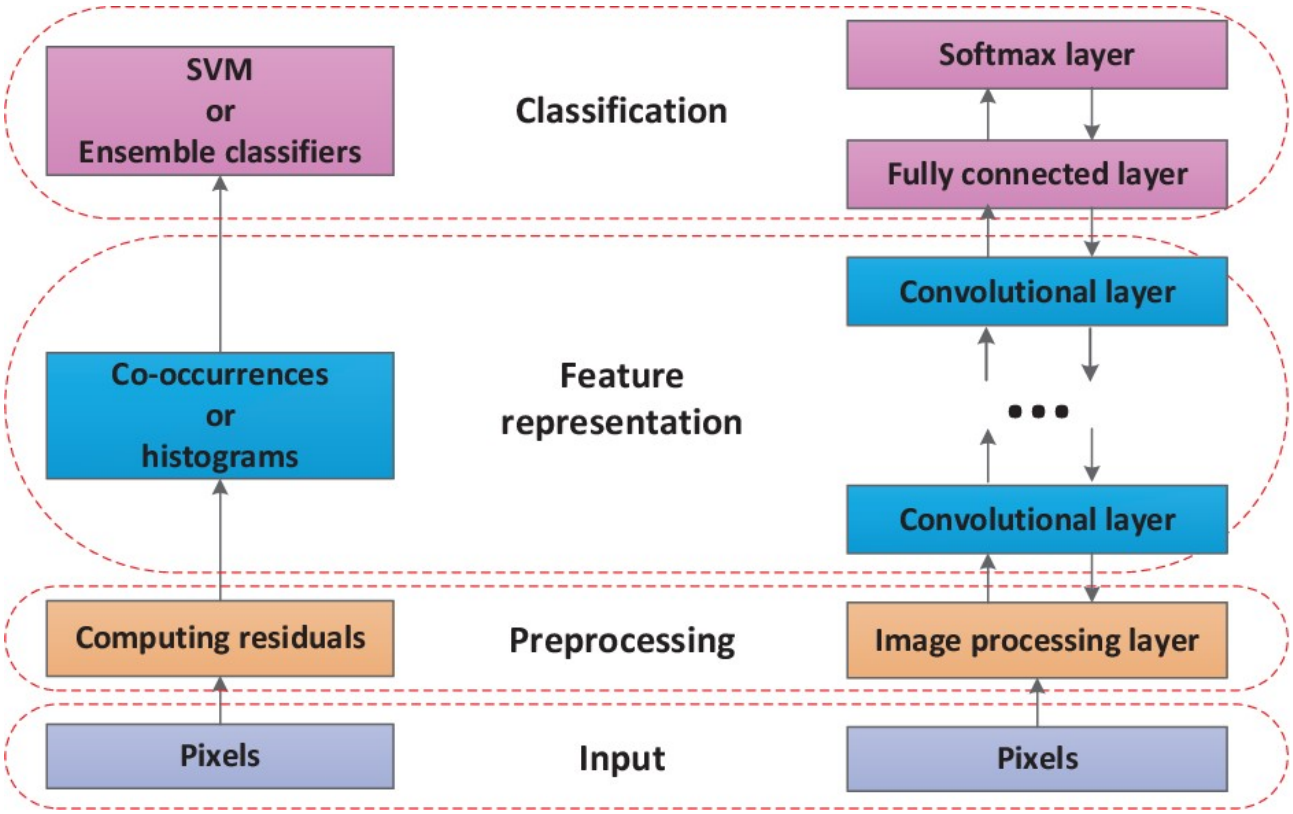
$$R_{(i,j)} = \theta(N_{ij}) - X_{(i,j)} \qquad (۱)$$

کرنل‌ها فقط شامل پیکسل‌هایی هستند که در جهت افقی/عمودی مرتب شده‌اند که از مدل‌های ثابت، خطی و درجه دوم موجود در تکه‌های تصویر محلی گرفته شده‌اند [۲۱]. SRM می‌تواند به عنوان یک سیستم استخراج ویژگی تک مرحله‌ای در نظر گرفته شود. اولین لایه‌ی آن یک بانک فیلتر است که با انواع مختلف آشکارسازهای لبه ساخته شده است.

از لحاظ ساختاری SRM به شبکه عصبی پیچشی شباهت دارد، تفاوت اصلی شبکه عصبی پیچشی و SRM این است که لایه اول SRM با انواع مختلف آشکارسازی‌های لبه پیاده‌سازی می‌شود ولی در شبکه عصبی پیچشی بانک‌های فیلتر با مقادیر تصادفی مقداردهی می‌شوند و به صورت با نظارت آموزش داده می‌شوند [۲۲].

۴ شبکه‌های عصبی پیچشی

شبکه‌های عصبی پیچشی عملکرد رضایت‌بخشی در پردازش داده‌های دو بعدی مانند صوت و ویدیو نشان داده‌اند. اولین نمونه موفق شبکه عصبی پیچشی در دهه ۹۰ میلادی بود که کار آن تشخیص اعداد و کاراکترهای دست نویس بود. در این بخش چهار مدل شبکه عصبی پیچشی در نهان‌کاوی معرفی می‌شود.



شکل ۳. سمت چپ (معماری سنتی شبکه عصبی برای نهان‌کاوی با استفاده از ویژگی‌های دست ساز) سمت راست (مدل پیشنهادی GNCNN برای نهان‌کاوی) [۲۳]

۱.۴ مدل GNCNN برای نهان‌کاوی

بین هونگ کیان [1] و همکاران [۲۳] در سال ۲۰۱۶ یک مدل پیشنهادی از شبکه عصبی پیچشی برای تشخیص پوشانه و گنجانه بودن تصاویر ارائه دادند. نهان‌کاوی کاملا متفاوت از وظایف هوش مصنوعی است. در واقع نویز گنجانه یک نوع سیگنال بسیار ضعیف است که توسط ادراک بشر درک نمی‌شود و این نویز در وظایف هوش مصنوعی هم نادیده گرفته می‌شود. در مدل پیشنهادی که در واقع یک مدل شبکه عصبی پیچشی سفارشی است این مشکل حل شده است و به آن GNCNN [2] گفته می‌شود.

شکل ۳ سمت راست، معماری مدل پیشنهادی را نشان می‌دهد که از ۳ لایه ساخته شده است و پیکسل‌ها را به عنوان ورودی می‌گیرد. این ۳ لایه عبارتند از : لایه پردازش تصویر، چندلایه پیچشی برای بازنمایی ویژگی‌ها، چند لایه تمام متصل برای طبقه‌بندی. در سمت چپ شکل ۳ معماری سنتی که استخراج ویژگی به صورت دستی انجام می‌شد و ویژگی‌ها به یک طبقه‌بند مانند ماشین بردار پشتیبان داده می‌شد، نشان داده شده است.

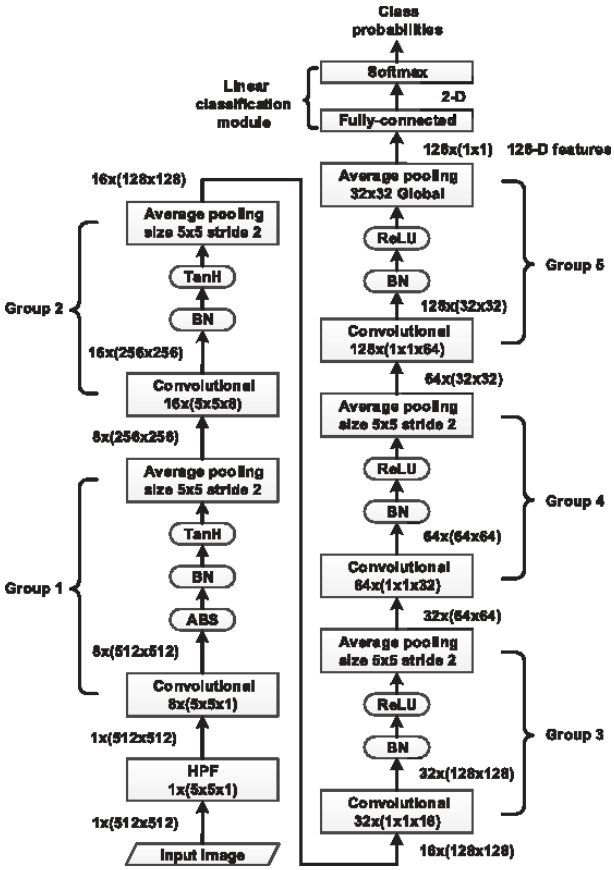
لایه پردازش تصویر: در این لایه یک عملیات فیلترینگ که در طول آموزش ثابت است انجام می‌شود. نویز گنجانه اضافه شده به تصویر پوشانه به شدت توسط محتوای تصویر فشرده شده است به خاطر همین برای کاهش این فشرده‌سازی و تقویت سیگنال گنجانه این فیلترینگ اعمال می‌شود.

لایه پیچشی: بعد از اعمال عملیات پیش پردازش برای تقویت سیگنال گنجانه در لایه‌ی پردازش تصویر، حال به صورت سلسله مراتبی پاسخ سیگنال گنجانه را از محلی به سراسری در لایه‌های پیچشی منتشر کردند. ورودی و خروجی هر لایه پیچشی مجموعه‌ای از آرایه‌ها هستند که به آن‌ها نگاشت ویژگی [3] گفته می‌شود. نگاشت ویژگی‌های ورودی وارد لایه پیچش شده و فیلتر روی آن انجام می‌شود (ضرب نقطه ای در کرنل) و خروجی آن از یک تابع غیرخطی گوسین به عنوان تابع فعال‌ساز عبور داده می‌شود.

لایه طبقه‌بند: ماژول طبقه بندی شامل چندین لایه تمام متصل است (تمام نورون‌های ورودی به تمام نورون‌های خروجی متصل است). ویژگی‌های یادگرفته شده به این لایه داده می‌شوند و در لایه بالاتر یک تابع غیرخطی فعال‌ساز به نام بیشینه هموار [4] برای تولید یک توزیع بر روی تمام برچسب‌های کلاس استفاده شده است. در واقع این تابع بزرگ‌ترین عدد را بزرگ‌ترین احتمال در نظر می‌گیرد، و در نهایت باید جمع احتمال‌ها یک شود.

[1] Yinlong Qian [2] Gaussian-Neuron CNN

[3] feature map [4] softmax



شکل ۴. ساختار شبکه عصبی پیچشی پیشنهادی Xu-NET [۲۴]

۲.۴ مدل شبکه عصبی پیچشی Xu-NET برای نهان‌کاوی

در این مدل گوانشو زو و همکاران [۲۴]، یک معماری ساختاری برای شبکه عصبی پیچشی در نهان‌کاوی ارائه دادند. در این معماری، مقادیر مطلق عناصر در نگاشت‌های ویژگی که از اولین لایه پیچشی تولید شده‌اند را برای تسهیل و بهبود مدل آماری در لایه‌های بعدی می‌گیرند. برای جلوگیری از پدیده بیش‌برازش [۱] دامنه داده‌ها را به واسطه تابع تانژانت هایپربولیک در لبه‌های شبکه محدود کردند و از پیچش‌های ۱ × ۱ در لایه‌های عمیق‌تر برای جلوگیری از افزایش قدرت مدل‌سازی استفاده کردند. شکل ۴ معماری کلی شبکه عصبی پیچشی را نشان می‌دهد. به منظور افزایش نسبت سیگنال به نویز از لایه HPF در نظر گرفته شده در مقاله هونگ کیان و همکاران [۲۳] استفاده شده که پارامترهای آن‌ها در طول آموزش بهینه نشده و این پارامترها در مراحل اولیه آموزش قرار می‌گیرد. این مدل هم به طور کلی از ماژول پیچشی برای استخراج ویژگی‌های ۱۲۸ بعدی از تصاویر و یک ماژول طبقه‌بندی (لایه‌های تمام متصل و تابع بیشینه هموار) که بردار ویژگی‌ها را به احتمال خروجی هر کلاس منتقل می‌کند، تشکیل شده است.

۳.۴ ساختار شبکه عصبی پیچشی پیشنهادی Xu-NET [۲۴]

CNN پیشنهادی جیان یه [۲] و همکاران [۶]، از ۱۰ لایه تشکیل شده است و با یک لایه تمام متصل و تابع بیشینه هموار خاتمه می‌یابد. بعد از هر لایه پیچشی یک تابع فعال ساز غیر خطی اجرا و عملیات کاهش بعد از لایه ۳ اعمال می‌شود. برخلاف سایر معماری‌های CNN که از دو یا چند لایه تمام متصل استفاده می‌کنند، در این شبکه فقط یک لایه دو طرفه تمام متصل به کار گرفته می‌شود.

این امر به این دلیل است که لایه‌های تمام متصل به پارامترهای زیادی وصل هستند که باید آموزش داده شوند و این به راحتی ممکن است منجر به پدیده بیش برازش شود، به خصوص زمانی که مجموعه آموزش به اندازه کافی بزرگ نباشد. عمق و عرض شبکه و اندازه فیلترها توسط آزمایشات و براساس مبادله بین عملکرد و پیچیدگی مدل تعیین می‌شود.

به طور خلاصه این معماری در ۳ مورد زیر خلاصه می‌شود:

(۱) اولین لایه CNN از ماژول پیش‌پردازش برای محاسبه میزان نویز استفاده می‌کند. در این مدل به جای استفاده از مقادیر تصادفی در وزن‌های لایه اول، از کرنل‌های SRM برای زودتر همگرا شدن شبکه استفاده می‌کنند.

(۲) در این شبکه در لایه‌های ابتدایی برای تقویت سیگنال‌های ناشی از نهان‌نگاری، از تابع فعال‌ساز جدیدی به نام TLU [۳] استفاده شده است. این تابع در فرمول (۲) آورده شده است.

$$f(x)=\begin{cases}-T & x<-T\\ x & -T<x<T\\ T & x>T\end{cases} \qquad (۲)$$

(۳) در نهایت عملکرد نهان‌کاوی را با انتخاب کانال نهان‌نگاری شده مشخص در مجموعه آموزشی شبکه عصبی پیچشی، ارتقا دادند.

۴.۴ مدل شبکه عصبی پیچشی YEDROUDJ-NET برای نهان‌کاوی مکانی

مهدی یادروج [۴] و همکاران [۲۵]، یک مدل از شبکه عصبی پیچشی را ارائه دادند که شمای کلی آن در شکل ۵ نشان داده شده است. این شبکه از یک بلوک پیش پردازشی، پنج بلوک پیچشی و یک بلوک تمام متصل ساخته شده که این بلوک خود از سه لایه کاملا متصل و به دنبال آن تابع بیشینه هموارساخته شده است. این شبکه یک توزیع احتمالی را در دو برچسب کلاس ایجاد می‌کند. بلوک پیش‌پردازش، تصویر پوشانه یا گنجانه ورودی را گرفته و آن را از یک فیلتر عبور می‌دهد. با عبور از این فیلتر اجزای سازنده نویز استخراج می‌شود. سپس تصویر پیش‌پردازش شده شبکه را تغذیه می‌کند. مشاهدات قبلی، دیدند که بدون این فیلتر شبکه عصبی پیچشی به آرامی همگرا می‌شود. این پیش‌پردازش، عمدتا محتوای تصویر را کاهش می‌دهد، دامنه حرکتی را محدود کرده در نتیجه

[1] overfitting

[2] Jian Ye [3] Truncated Linear Unit [4] Mehdi Yedroudj

نسبت سیگنال به نویز بین سیگنال ضعیف گنجانه (در صورت حضور) و سیگنال تصویر را افزایش می‌دهد. این کار منجر به یادگیری بهتر شبکه عصبی پیچشی روی یک سیگنال فشرده و مقاوم‌تر می‌شود. در این طرح، اندازه تمام کرنل‌ها ۵ × ۵ در نظر گرفته شده که بخش مرکزی آن با عناصر SRM آغاز شده و باقی عناصر با مقدار صفر لایه‌گذاری می‌شوند. هیچ نرمال‌سازی‌ای برای مقادیر کرنل‌ها انجام نمی‌شود. به طور مشابه با Xu-Net، ماژول‌های پیچشی از ۵ بلوک (از بلوک ۱ تا بلوک ۵) تشکیل شده‌اند که از آن‌ها برای استخراج ویژگی از اختلاف موجود بین تصاویر پوشانه و گنجانه استفاده می‌شود.

۵ تحلیل نتایج به دست آمده برای چهار مدل معماری شبکه عصبی پیچشی

هدف اصلی در این بخش مقایسه عملکرد چهار مدل شبکه عصبی پیچشی معرفی شده شامل GNCNN [۲۳] و Xu-Net [۲۴] و Yedrouj-Net [۲۵] و Ye-Net [۶] و روش SRM [۲۱] به عنوان یک روش معمولی و سنتی است. این مقایسه می‌تواند به محققان این حوزه کمک کند تا شناخت مناسبی از کارایی فعلی روش‌های شبکه عصبی پیچشی نسبت به یکدیگر و نسبت به روش‌های سنتی داشته باشند.

در جدول ۱، خطای احتمالی پنج روش نهان‌کاوی مورد نظر برای کشف الگوریتم های جاسازی WOW و S-UNIWARD برای درصدجاسازی‌های ۰٫۲، ۰٫۳، ۰٫۴ و ۰٫۵ نشان داده شده است. پایگاه داده مورد آزمایش برای پنج تکنیک نهان‌کاوی BOSSbase در نظر گرفته شده که شامل ۱۰۰۰۰ تصویر ۵۱۲ × ۵۱۲ است. برای سادگی محاسبات در این پنج روش اندازه تصاویر را به ابعاد ۲۵۶ × ۲۵۶ تغییر دادند.

در درصد جاسازی‌های ۰٫۳ و ۰٫۵ مدل GNCNN که می‌توان گفت یکی از اولین مدل‌های شبکه عصبی پیچشی به شمار می‌رود، خطای احتمالی به دست آمده برای هر دو الگوریتم به میزان ۳٪ تا ۴٪ بیشتر از مدل SRM می‌باشد. مدتی بعد مدل‌های دیگری مانند Xu-Net، Ye-Net و Yedrouj-Net معرفی شدند که هدف آن‌ها کاهش این مقدار خطا با استفاده از بهبود مدل شبکه عصبی پیچشی بود.

برای الگوریتم WOW، مدل Yedrouj-Net برای درصدجاسازی‌های ۰٫۲ و ۰٫۴ به ترتیب به اندازه ۸٪ و ۱۱٪ خطای احتمالی کمتری نسبت به مدل SRM دارد. همچنین این مدل برای الگوریتم S-UNIWARD به خطای احتمالی برابری در درصد جاسازی ۰٫۲ و ۲٪ کمتر از روش نهان‌کاوی SRM در درصد جاسازی ۰٫۴ رسیده است. این نشان می‌دهد که با استفاده از الگوریتم یادگیری عمیق توانستند به درصد خطای قابل قبولی دست پیدا کنند.

اما برای مقایسه سایر مدل‌های شبکه عصبی پیچشی مشاهده می‌شود که در درصد جاسازی ۰٫۴، مدل Yedrouj-Net برای الگوریتم WOW به مراتب خطای احتمالی بهتری به میزان ۶٪ تا ۱۵٪ نسبت به سایر روش‌ها کسب کرده است. هم‌چنین این مدل برای الگوریتم S-UNIWARD هم به میزان ۴٪ تا ۸٪ خطای احتمالی کمتری نسبت به سایر مدل‌ها به دست آورده است.

در نهایت همانطورکه در جدول ۱ مشاهده می‌شود، در میان این مدل‌ها تا کنون مدل Yedrouj-Net عملکرد بهتری از خود نشان داده است.

ذکر این نکته الزامی است که مدل Xu-Net همیشه برتر از مدل SRM نیست. برای شکست این مدل باید یا از شبکه عصبی پیچشی گروهی استفاده کرد و یا اندازه داده‌های یادگیری را افزایش داد.

۶ نتیجه

تا کنون روش‌های نهان‌کاوی تصاویر دیجیتالی ابتدا روی ویژگی‌های دست ساز پیچیده متمرکز بوده‌اند. به همین جهت، ابتدا ویژگی‌ها به صورت دستی استخراج شده و سپس برای آموزش به طبقه‌بند جهت تشخیص گنجانه یا پوشانه بودن تصاویر می‌دادند. امروزه با استفاده از مدل‌های یادگیری عمیق می‌توان ویژگی‌ها را از تصاویر با ابعاد بالا به صورت خودکار استخراج کرد و طبقه‌بند را در جهت به دست آوردن ویژگی‌های دقیق‌تر آموزش داد. به عبارت دیگر مراحل استخراج ویژگی و طبقه‌بندی تحت یک معماری واحد قرار گرفتند.

مدل‌های جدید نهان‌کاوی بر پایه استخراج خودکار ویژگی‌های با ابعاد بالا، توسط الگوریتم شبکه عصبی پیچشی پیاده‌سازی می‌شوند که در این مقاله به معرفی چهار مدل GNCNN، Yedrouj-Net، Xu-Net و Ye-Net پرداخته شد. در پایان طبق تحلیل نتایج به دست آمده برای الگوریتم‌های نهان‌نگاری WOW و S-UNIWARD، مشاهده شد که خطای احتمالی مدل Yedrouj-Net در درصد جاسازی‌های ۰٫۲ و ۰٫۴ کمتر و یا برابر با مدل SRM است.

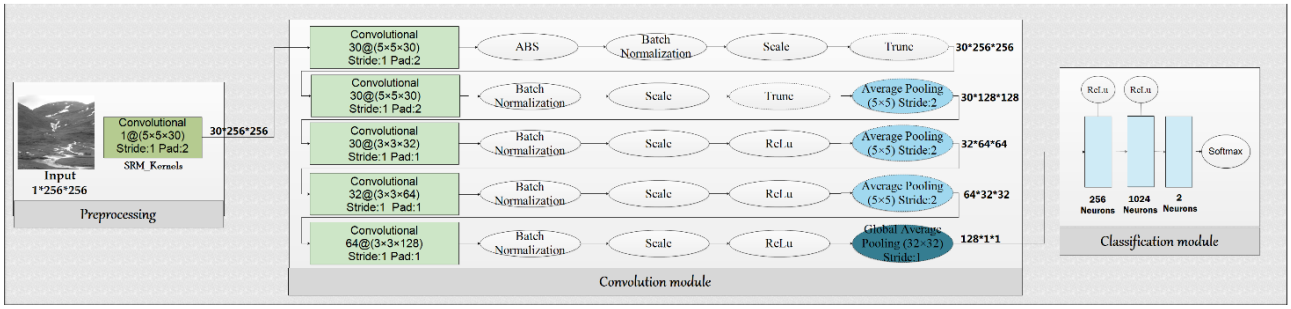
مراجع

[1] Yambem Jina Chanu, Themrichon Tuithung, and Kh Manglem Singh. A short survey on image steganography and steganalysis techniques. In *2012 3rd National Conference on Emerging Trends and Applications in Computer Science*, pages 52–55. IEEE, 2012.

[2] Alaa A Jabbar Altaay, Shahrin Bin Sahib, and Mazdak Zamani. An introduction to image steganography techniques. In *2012 International Conference on Advanced Computer Science Applications and Technologies (ACSAT)*, pages 122–126. IEEE, 2012.

[3] A Rutherford. Steganography in lossy and lossless images. https://github.com/pepper-project/pequin, 2014.

[4] Vojtěch Holub, Jessica Fridrich, and Tomáš Denemark.



شکل ۵. معماری شبکه عصبی پیچشی Yedroudj-Net [۲۵]

جدول ۱. خطای احتمالی به دست آمده برای پنج روش نهان‌کاوی

	IBOSSbase 256×256							
	WOW				S-UNIWARD			
Steganalysis / Payload	0.2 bpp	0.3 bpp	0.4 bpp	0.5 bpp	0.2 bpp	0.3 bpp	0.4 bpp	0.5 bpp
SRM [21]	36.5%	31.2%	25.5%	22.1%	36.6%	31.5%	24.7%	21.4%
GNCNN [23]	-	34.3%	29.3%	24.8%	-	35.6%	30.9%	26.3%
Yedrouj-Net [25]	27.8%	-	14.1%	-	36.7%	-	22.8%	-
Xu-Net [24]	32.4%	-	20.7%	-	39.1%	-	27.2%	-
Ye-Net [6]	33.1%	-	23.2%	-	40.0%	-	31.2%	-

Entropy feature based on 2d gabor wavelets for jpeg steganalysis. In *International Conference on Security, Privacy and Anonymity in Computation, Communication and Storage*, pages 59–72. Springer, 2016.

[11] GM Deepa, R Keerthi, N Meghana, and K Manikantan. Face recognition using spectrum-based feature extraction. *Applied Soft Computing*, 12(9):2913–2923, 2012.

[12] Xiaofeng Song, Fenlin Liu, Xiangyang Luo, Jicang Lu, and Yi Zhang. Steganalysis of perturbed quantization steganography based on the enhanced histogram features. *Multimedia Tools and Applications*, 74(24):11045–11071, 2015.

[13] Xiaohua Xie, Jianhuang Lai, and Wei-Shi Zheng. Extraction of illumination invariant facial features from a single image using nonsubsampled contourlet transform. *Pattern Recognition*, 43(12):4177–4189, 2010.

[14] Jan Kodovsky, Jessica Fridrich, and Vojtěch Holub. Ensemble classifiers for steganalysis of digital media. *IEEE Transactions on Information Forensics and Security*, 7(2):432–444, 2011.

[15] Nan Liu and Han Wang. Ensemble based extreme learning machine. *IEEE Signal Processing Letters*, 17(8):754–757, 2010.

Universal distortion function for steganography in an arbitrary domain. *EURASIP Journal on Information Security*, 2014(1):1–13, 2014.

[5] Vojtěch Holub and Jessica Fridrich. Designing steganographic distortion using directional filters. In *2012 IEEE International workshop on information forensics and security (WIFS)*, pages 234–239. IEEE, 2012.

[6] Jian Ye, Jiangqun Ni, and Yang Yi. Deep learning hierarchical representations for image steganalysis. *IEEE Transactions on Information Forensics and Security*, 12(11):2545–2557, 2017.

[7] John Babu, Sridevi Rangu, Pradyusha Manogna, et al. A survey on different feature extraction and classification techniques used in image steganalysis. *Journal of Information Security*, 8(03):186, 2017.

[8] Farshid Farhat and Shahrokh Ghaemmaghami. Towards blind detection of low-rate spatial embedding in image steganalysis. *IET Image Processing*, 9(1):31–42, 2015.

[9] Hassan Karimi, Mahrokh G Shayesteh, and Mohammad Ali Akhaee. Steganalysis of jpeg images using enhanced neighbouring joint density features. *IET Image Processing*, 9(7):545–552, 2015.

[10] Xiaofeng Song, Zhiyuan Li, Liju Chen, and Jiong Liu.

[16] Guang-Bin Huang, Xiaojian Ding, and Hongming Zhou. Optimization method based extreme learning machine for classification. *Neurocomputing*, 74(1-3):155–163, 2010.

[17] Yakoub Bazi, Naif Alajlan, Farid Melgani, Haikel Al-Hichri, Salim Malek, and Ronald R Yager. Differential evolution extreme learning machine for the classification of hyperspectral images. *IEEE Geoscience and Remote Sensing Letters*, 11(6):1066–1070, 2013.

[18] Vasily Sachnev, Savitha Ramasamy, Suresh Sundaram, Hyoung Joong Kim, and Hee Joon Hwang. A cognitive ensemble of extreme learning machines for steganalysis based on risk-sensitive hinge loss function. *Cognitive Computation*, 7(1):103–110, 2015.

[19] Peiqing Liu, Fenlin Liu, Chunfang Yang, and Xiaofeng Song. Improving steganalysis by fusing svm classifiers for jpeg images. In *2015 International Conference on Computer Science and Mechanical Automation (CSMA)*, pages 185–190. IEEE, 2015.

[20] Fengyong Li, Xinpeng Zhang, Bin Chen, and Guorui Feng. Jpeg steganalysis with high-dimensional features and bayesian ensemble classifier. *IEEE signal processing letters*, 20(3):233–236, 2013.

[21] Jessica Fridrich and Jan Kodovsky. Rich models for steganalysis of digital images. *IEEE Transactions on information Forensics and Security*, 7(3):868–882, 2012.

[22] Shunquan Tan and Bin Li. Stacked convolutional auto-encoders for steganalysis of digital images. In *Signal and information processing association annual summit and conference (APSIPA), 2014 Asia-Pacific*, pages 1–4. IEEE, 2014.

[23] Yinlong Qian, Jing Dong, Wei Wang, and Tieniu Tan. Deep learning for steganalysis via convolutional neural networks. In *Media Watermarking, Security, and Forensics 2015*, volume 9409, pages 171–180. SPIE, 2015.

[24] Guanshuo Xu, Han-Zhou Wu, and Yun-Qing Shi. Structural design of convolutional neural networks for steganalysis. *IEEE Signal Processing Letters*, 23(5):708–712, 2016.

[25] Mehdi Yedroudj, Frédéric Comby, and Marc Chaumont. Yedroudj-net: An efficient cnn for spatial steganalysis. In *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 2092–2096. IEEE, 2018.