پنهان شکنی روش پنهان‌نگاری بر مبنای اختلاف مقادیر پیکسل‌ها با بازه‌بندی تصادفی با استفاده از شبکه عصبی

وجیهه ثابتی
دانشگاه صنعتی اصفهان
Vajiheh.sabeti@gmail.com

شادرخ سماوی
دانشگاه صنعتی اصفهان
Samavi96@cc.iut.ac.ir

محمدرضا احمدزاده
دانشگاه صنعتی اصفهان
Ahmadzadeh@cc.iut.ac.ir

چکیده: در این مقاله روشی برای پنهان‌شکنی یکی از روش‌های پنهان‌نگاری برمبنای اختلاف مقادیر پیکسل‌ها ارائه شده است و نشان داده شده است علی‌رغم ادعای ارائه کنندگان روش مزبور مبنی بر امن بودن روش فوق این روش قابل شکست بوده است. این روش پنهان-نگاری نیز مانند دیگر روش‌های پنهان‌نگاری برمبنای اختلاف مقادیر پیکسل‌ها، جاسازی را در مقدار تفاوت پیکسل‌های بلاک‌های دوتایی انجام می‌دهد. بنابراین هیستوگرام تفاوت پیکسل‌ها در تصویر حاصل از این روش دارای توزیعی متفاوت از تصاویر نرمال است. برای تشخیص و تمایز این دو گروه از تصاویر، تعدادی پارامتر از هیستوگرام تفاوت تصویر استخراج شده است که در اثر جاسازی به طور محسوسی تغییر می-کنند. یک شبکه عصبی پرسپترون چندلایه برای رده بندی این تصاویر با استفاده از پارامترهای استخراج شده پیشنهاد شده است. در پیاده سازی انجام شده شبکه پیشنهادی با دقت ۹۸/۶٪ موفق به رده بندی تصاویر مجموعه تست شده است.

واژه های کلیدی: پنهان‌نگاری، پنهان‌شکنی، شبکه عصبی، pixel_value difference

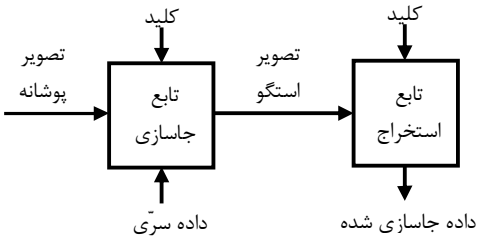
۱- مقدمه

پنهان‌نگاری هنر ارتباط پنهانی است و هدف آن پنهان کردن ارتباط به وسیله قرار دادن پیام در یک رسانه پوششی است به گونه‌ای که کمترین تغییر قابل کشف را در آن ایجاد نماید و نتوان موجودیت پیام پنهان در رسانه را حتی به صورت احتمالی آشکار ساخت[۱].

تفاوت اصلی رمزنگاری و پنهان‌نگاری آن است که در رمزنگاری هدف اختفاء محتویات پیام است و نه به طور کلی وجود پیام، اما در پنهان-نگاری هدف مخفی کردن هر گونه نشانه‌ای از وجود پیام است. در مواردی که تبادل اطلاعات رمز شده مشکل آفرین است باید وجود ارتباط پنهان گردد[۲].

به صورت کلی در سیستم‌های اختفاء اطلاعات سه عنصر اصلی ظرفیت، امنیت و مقاومت دخیل هستند. در روش های پنهان‌نگاری، ظرفیت و امنیت اهمیت اصلی را دارند.

تصاویر یکی از مهمترین رسانه‌های مورد استفاده در اینترنت هستند. از آنجایی که درک انسان از تغییرات در تصاویر محدود است، بنابراین تصاویر به عنوان نوعی رسانه پوششی مناسب در پنهان‌نگاری محسوب می‌شوند و الگوریتم‌های پنهان‌نگاری متعددی برای ساختارهای مختلف تصاویر ارائه شده است. در زمینه پنهان‌نگاری تصاویر، بعضی از اصطلاحات رایج شده‌اند. شکل ۱، اصطلاحات معمول و ارتباط بین آن‌ها را نشان می‌دهد[۳].



شکل۱: مدل کلی سیستم‌های پنهان‌نگاری در تصاویر

تصویر پوشانه[1]، تصویری است که در آن جاسازی انجام می‌شود و تصویر استگو[2] خروجی فرآیند جاسازی است. داده جاسازی شده، همان پیام سری موردنظر است. برای افزایش امنیت معمولاً از یک کلید برای جاسازی استفاده می‌شود.

با رواج یافتن روش‌های پنهان‌نگاری، علم دیگری با عنوان پنهان‌شکنی رونق یافت. پنهان‌شکنی، روش‌هایی برای حمله و شکست الگوریتم‌های پنهان‌نگاری است. برای یک ارتباط سری، تنها کشف و اثبات وجود یک داده مخفی در تصویر استگو، حمله موفق به حساب می‌آید. روش‌های پنهان‌شکنی از تغییرات ویژگی‌های آماری تصویر استفاده می‌کنند که در اثر استفاده از پنهان‌نگاری ایجاد شده‌اند[۴].

روش‌های پنهان‌نگاری داده در تصویر از دو شیوه خاص پیروی می‌کنند. گروهی از روش‌ها، از حوزه مکانی تصویر برای جاسازی پیام استفاده

[1] Cover image
[2] Stego image

می‌کنند. در این روش‌ها با تغییر اطلاعات پیکسل‌ها، برای مثال تغییر بیت‌هایی از مقدار پیکسل، داده را در تصویر مخفی می‌کنند. دسته دوم روش‌ها از حوزه تبدیل تصویر استفاده می‌کنند. در این روش‌ها، ابتدا تصویر با استفاده از یک تبدیل خاص مانند DCT یا DWT به حوزه تبدیل برده می‌شود. سپس با تغییر مقادیر ضرایب تبدیل، جاسازی پیام انجام می‌شود[۵].

روش‌های جاسازی در حوزه مکان، روش‌های ساده‌تر و متداول‌تری هستند. استفاده از بیت‌های کم ارزش پیکسل‌های تصویر از روش‌های معمول در پنهان‌نگاری است. دو گروه عمده از این روش‌ها LSB-F [3] و LSB-M [4] هستند. در روش LSB-F داده موردنظر مستقیماً در بیت کم ارزش قرار داده می‌شود. اما در روش LSB-M، در صورت عدم تطابق بیت کم ارزش با داده موردنظر، مقدار پیکسل به صورت تصادفی یک واحد افزایش یا کاهش داده می‌شود. ظرفیت مناسب و عدم حساسیت چشم به تغییرات بیت کم ارزش از مزایای این دسته از روش‌ها است[۶].

تا به حال روش‌هایی برای شکست الگوریتم‌های LSB-F و LSB-M ارائه شده است. ضعف اصلی روش LSB-F، تولید زوج مقدارها[5] در هیستوگرام تصویر استگو است و تا به حال روش‌های مختلفی مانند χ^2، RS و ... از این واقعیت برای کشف این روش استفاده کرده‌اند[۷و۸]. اگرچه روش LSB-M این نقطه ضعف را ندارد، اما حملات موفقی به این روش نیز تا به حال ارائه شده است که درصد دقت متفاوتی دارند[۶و۹].

Wu و همکارانش در [۱۰] روش جدیدی برای جاسازی در حوزه مکان تصویر پیشنهاد کرده‌اند که به جای جاسازی در بیت‌های کم ارزش پیکسل‌ها، جاسازی را در مقدار تفاوت روشنایی پیکسل‌ها (PVD)[6] انجام می‌دهد. در این روش تصویر به بلاک‌هایی مجزا از دو پیکسل متوالی تقسیم بندی می‌شود. جاسازی اطلاعات در مقدار تفاوت پیکسل‌های درون بلاک انجام می‌شود. برای جاسازی یک زیررشته از پیام، مقدار تفاوت با یک مقدار جدید جایگزین می‌شود. ظرفیت جاسازی هر بلاک به مقدار تفاوت در آن بلاک بستگی دارد. ارائه دهندگان این روش ادعا می‌کنند که چون جاسازی داده به صورت مستقیم در فضای پیکسلی تصویر انجام نمی‌شود، این روش در برابر حملاتی که به دنبال تغییرات در فضای پیکسلی تصویر[۸] و هیستوگرام حاصل از آن هستند[۷]، مقاوم است.

در[۱۱]، روش دیگری بر مبنای روش PVD ارائه شده است که ادعا می‌شود با حفظ امنیت روش PVD و بعلاوه حفظ کیفیت تصویر استگو، ظرفیت جاسازی در آن افزایش یافته است. در این روش برای جاسازی در بلاک‌های یکنواخت از روش LSB و برای جاسازی در بلاک‌های لبه از روش PVD استفاده می‌شود.

اما علی رغم ادعای ارائه دهندگان دو روش [۱۰] و [۱۱]، مبنی بر شکست ناپذیری در برابر حملات موجود، در [۱۲] ثابتی و همکارانش حمله ای برای این دو روش ارائه کردند.

Zhang و همکارش در [۱۳] روش دیگری برمبنای PVD پیشنهاد کرده‌اند که در آن توانسته‌اند تا حدی بر نقطه ضعف اصلی روش PVD فائق آیند. تا به حال روشی برای کشف این روش ارائه نشده است. از این روش در ادامه با عنوان MPVD [7] نام برده می‌شود و سعی می‌شود روشی برای شکست آن پیشنهاد شود. در حمله‌ای که در این مقاله ارائه می‌شود، از شبکه عصبی به عنوان یک رده بند استفاده خواهد شد.

معمولاً روش‌های پنهان شکنی کور[8] از رده بند استفاده می کنند. چنین روش‌هایی یک رده بند طراحی می‌کنند که براساس مجموعه آموزشی از اشیاء پوشانه و استگو حاصل تعدادی از الگوریتم‌های پنهان‌نگاری، آموزش داده می‌شود. طبقه بندی براساس آن دسته از ویژگی‌های ذاتی تصاویر طبیعی انجام می‌شود که امکان دارد بعد از گذراندن فرآیند جاسازی نقض شوند. روش پنهان‌شکنی کور براساس آمارهای درجه بالا [۱۴] و روش پنهان شکنی کور تصاویر JPEG [۱۵] دو نمونه از این روش‌ها هستند. هدف روش‌های پنهان‌شکنی کور، کشف تعدادی از الگوریتم‌های پنهان نگاری است. روش‌های پنهان‌شکنی مخصوص یک الگوریتم منحصربه فرد نیز وجود دارند که از رده‌بندها استفاده می‌کنند. به عنوان نمونه می‌توان به استفاده از این مفهوم در [۱۶] برای کشف روش LSB-M اشاره کرد.

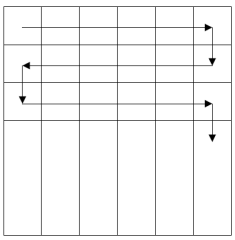
طراحی یک رده‌بند، شامل دو گام می‌باشد. گام ابتدایی، پیدا کردن و محاسبه ویژگی‌هایی است که قادر به تسخیر تغییرات آماری ناشی از فرآیند جاسازی در تصویر باشند. گام دوم، انتخاب یک رده‌بند است که تفاوت‌های تسخیر شده بوسیله ویژگی‌ها را بیشترین مقدار کند و دقت طبقه‌بندی زیادی داشته باشد[۵]. در مقاله ویژگی‌ها و رده‌بند استفاده شده برای حمله معرفی خواهد شد.

در ادامه مقاله روش PVD و MPVD به صورت مختصر معرفی می‌شوند. در بخش سوم تغییرات آماری ایجاد شده توسط روش MPVD مورد بحث قرار می‌گیرد. روش حمله‌ی پیشنهادی و نتایج عملی اعمال حمله بر روی تعدادی از عکس‌های نمونه در بخش‌های چهارم و پنجم ارائه می‌شود. نتیجه گیری نیز در بخش ششم آورده شده است.

[3] LSB Flipping (LSB-F)
[4] LSB Matching (LSB-M)
[5] Pair Of Value(POV)
[6] Pixel_Value Differencing(PVD)
[7] Modified PVD (MPVD)
[8] Blind Steganalysis

۲-معرفی روش های جاسازی PVD و MPVD

ابتدا روش PVD را بررسی می‌کنیم. در این روش، تصاویر به صورت ۸ بیتی با سطوح خاکستری فرض می‌شوند. تصویر رنگی را می‌توان به سه جزء خاکستری تجزیه کرد. ابتدا تصویر پوشانه به بلاک‌های دو پیکسلی تقسیم‌بندی می شوند. برای انجام این تقسیم‌بندی، تمام سطرهای تصویر به صورت زیگزاگ پیمایش می‌شوند. در شکل ۲ نحوه پیمایش زیگزاگی تصویر نشان داده شده است. فرض کنید بلاکی از دو پیکسل همسایه با مقادیر خاکستری g_i و g_{i+1} انتخاب شده است. مقدار تفاوت d در این بلاک برابر $g_{i+1}-g_i$ است که در بازه [۲۵۵و-۲۵۵] قرار دارد. نزدیک شدن d به 0 نشان دهنده بلاک یکنواخت و دور شدن d از 0 نشان دهنده بلاک لبه (غیریکنواخت) است.



شکل۲: چگونگی پیمایش زیگزاگی تصویر

در این روش اطلاعات در مقدار d هر بلاک جاسازی می‌شود. میزان اطلاعات قابل جاسازی در هر بلاک به مقدار d در آن بستگی دارد. بدین منظور اندازه d در n بازه دسته‌بندی می‌شود. به هر بازه یک اندیس(k) نسبت داده می‌شود و حد بالا و پایین هر بازه نیز به ترتیب u_k و l_k نامیده می‌شود. طول این بازه برابر با $w_k = u_k - l_k + 1$ است که باید توانی از ۲ انتخاب شود. برای مثال، می‌توان مقادیر اختلاف را در ۶ بازه با طول‌های ۱۲۸،۶۴،۳۲،۱۶،۸،۸ تقسیم‌بندی کرد. ظرفیت هر بازه به طول آن بستگی دارد و برای بازه با اندیس k، به صورت زیر $n = \log w_k$ تعریف می‌شود.

فرض کنید یک بلاک دو پیکسلی با مقدار تفاوت d انتخاب شده است که مقدار d به بازه k تعلق دارد و با توجه به طول این بازه، ظرفیت جاسازی n بیت داده را دارد. برای جاسازی بیت‌های پیام باید مقدار تفاوت بلاک با یک مقدار تفاوت جدید جایگزین شود. برای جاسازی n بیت از رشته بیتی پیام که مقدار b را دارد، مقدار d جدید (d') برای بلاک به صورت زیر محاسبه می‌شود:

$$d' = \begin{cases} l_k + b & d \geq 0 \\ -(l_k + b) & d < 0 \end{cases} \qquad (۱)$$

با توجه به این که b در بازه $[0, u_k - l_k]$ قرار دارد، بنابراین مقدار d' نیز در بازه $[l_k, u_k]$ باقی می ماند. این ویژگی باعث می‌شود تا استخراج پیام جاسازی شده از تصویر استگو امکان پذیر باشد.

بعد از محاسبه مقدار d'، با استفاده از تابع زیر مقادیر جدید پیکسل-های بلاک (g'_i, g'_{i+1}) در تصویر استگو محاسبه می‌شود:

$$(g'_i, g'_{i+1}) = f((g_i, g_{i+1}), m) = \qquad (۲)$$
$$\begin{cases} (g_i - \lceil m/2 \rceil, g_{i+1} + \lfloor m/2 \rfloor) & if\ d\ is\ odd \\ (g_i - \lfloor m/2 \rfloor, g_{i+1} + \lceil m/2 \rceil) & if\ d\ is\ even \end{cases}$$

که در آن m برابر $d' - d$ است. بلاک‌هایی قابلیت جاسازی دارند که مقدار پیکسل‌های جدید آن‌ها، که در رابطه بالا محاسبه می‌شود، در بازه [۲۵۵و۰] باقی بماند.

برای افزایش امنیت روش PVD، یک نسخه تغییر یافته از این روش پیشنهاد شده است که با عنوان MPVD از آن یاد می‌کنیم. در این روش برخلاف روش PVD، به جای استفاده از بازه‌های ثابت، از بازه‌های متغیر استفاده شده است. به عبارت دیگر، بازه‌های متناظر با بلاک‌های مختلف برحسب یک کلید $\beta \in [0,1]$ تولید می شوند. این پارامتر به صورت شبه تصادفی برای هر بلاک انتخاب می‌شود و حدود بالا و پایین بازها بدین صورت محاسبه می‌شوند:

$$l'_k = l_k + \lfloor \beta . w_k \rfloor \qquad (۳)$$
$$u'_k = u_k + \lfloor \beta . w_{k+1} \rfloor$$

K، اندیس بازه است. اگر $l'_k \leq |d| \leq u'_k\ (k > 1)$ باشد، به تعداد $\log_2(w_k)$ بیت در بلاک مورد نظر جاسازی می شود. اگر مقدار دهدهی بیت های انتخاب شده را b بنامیم، مقدار تفاوت جدید (d') بدین طریق قابل محاسبه است:

$$d' = \begin{cases} \underset{l'_k \leq e \leq u'_k, \bmod(e, w_k) = b}{\operatorname{argmin}} \left(|e - d|\right) & if\ d > 0 \\ -\left[\underset{l'_k \leq e \leq u'_k, \bmod(e, w_k) = b}{\operatorname{argmin}} \left(|e - d|\right) \right] & if\ d < 0 \end{cases} \qquad (۴)$$

به عبارت دیگر، d' نزدیکترین مقدار به d در میان مقادیر همان بازه است که باقیمانده b را در تقسیم به w_k داشته باشند. اگر $0 \leq |d| \leq u'_0$ باشد، d' اینگونه محاسبه می‌شود:

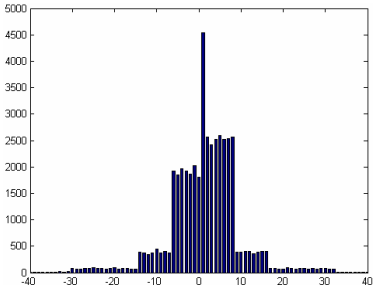
$$d' = \underset{-u'_0 \leq e \leq u'_0, \bmod(e, w_0) = b}{\arg\min} \left(|e - d|\right) \qquad (۵)$$

پس از محاسبه d'، مقدار دو پیکسل در تصویر استگو با استفاده از رابطه (۲) محاسبه می‌شود. گیرنده با استفاده از روش زیر ،داده جاسازی شده را استخراج می کند.

$$b=\begin{cases}\mathrm{mod}(d', w_0) & if\ 0\le|d'|\le u'_0\\ \mathrm{mod}(d', w_k) & if\ l'_k\le|d'|\le u'_k\ (k>0)\end{cases} \quad (۶)$$

۳- تغییرات آماری حاصل از روش MPVD

علیرغم ادعای ارائه‌دهندگان روش PVD ساده مبنی بر امنیت این روش در برابر حملات شناخته شده موجود، تا به حال دو روش حمله موفق مختلف [۱۲و۱۶] برای این روش جاسازی ارائه شده است. در هر دو روش حمله، از آنالیز هیستوگرام استفاده شده است. با توجه به اینکه جاسازی در روش PVD به صورت مستقیم در پیکسل‌های تصویر انجام نمی‌شود، بنابراین در هیستوگرام اصلی تصویر تغییر محسوس و قابل کشفی رخ نمی‌دهد. در این روش داده مورد نظر در مقدار تفاوت پیکسل‌ها جاسازی می‌شود، بنابراین کلید اصلی حمله به این روش جاسازی، بررسی هیستوگرام تفاوت جفت پیکسل های درون بلاک‌های دوتایی است. برای مثال در شکل ۳، هیستوگرام تفاوت تصویر Lena بعد از جاسازی ۱۰۰٪ به روش PVD ساده نشان داده شده است. به وضوح وجود گام های غیر عادی در این هیستوگرام دیده می شود.



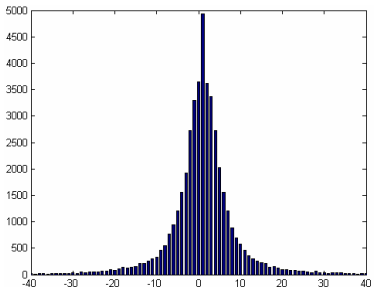
شکل۳: هیستوگرام تفاوت تصویر Lena بعد از اعمال روش PVD

در روش PVD ساده طول بازه ها (w_k) برای تمام بلاک‌ها ثابت است. در این روش به گونه‌ای مقدار تفاوت جدید (d') محاسبه می‌شود که داده مورد نظر از رابطه $d' \bmod w_k$ قابل استخراج باشد. با توجه به اینکه در این روش طول بازه همیشه توانی از ۲ انتخاب می‌شود، بنابراین نتیجه استخراج شده از d' ، شامل چند بیت از LSB آن می باشد که تعداد آن به مقدار w_k بستگی دارد. بنابراین در روش PVD ساده، جاسازی در LSB های مقدار تفاوت جفت پیکسل های درون بلاک‌ها انجام می‌شود. به عبارت دیگر می‌توان روش PVD ساده را همانند روش LSB-F تفسیر کرد. بدین ترتیب، گام‌های غیر عادی در هیستوگرام به دلیل رویداد جفت شدن در جاسازی به روش LSB-F رخ داده است. از این تفسیر در حمله [۱۲] استفاده شده است و با اعمال حمله χ^2 بر روی تصویر حاصل از تفاوت پیکسل‌ها، وجود داده جاسازی شده به روش PVD ساده کشف شده است.

ارائه دهندگان حمله [۱۳]، پس از استفاده از گام‌های غیر عادی در هیستوگرام تفاوت برای شکستن روش PVD ساده، روش MPVD را پیشنهاد کرده‌اند و ادعا می‌کنند این روش از ایجاد گام‌های غیرعادی جلوگیری می‌کند. تنها تفاوت روش جدید با روش اصلی این است که در روش MPVD طول بازه‌ها به صورت شبه تصادفی و متغیر برای هر بلاک انتخاب می‌شوند و الزامی برای توانی از ۲ بودن طول بازه‌ها وجود ندارد. بنابراین استخراج داده جاسازی شده از رابطه بالا نشان می‌دهد که در روش MPVD جاسازی الزاماً در LSB مقدار تفاوت انجام نمی‌شود. بدین ترتیب امکان رخ دادن جفت‌ها در این روش وجود ندارد.

تجربه نشان می‌دهد که هیستوگرام تفاوت، برای تصاویر عادی و بدون داده جاسازی تقریباً توزیعی مشابه با توزیع تعمیم یافته گوسی دارد[۱۷]. در شکل ۴، یک نمونه از هیستوگرام تفاوت تصویر Lena نشان داده شده است که توزیع شبیه گوسین دارد.

انتظار می‌رود که جاسازی به روش MPVD در مقادیر تفاوت باعث شود که این هیستوگرام در تصویر استگو توزیع گوسین مناسبی نداشته باشد. به عبارت دیگر، هیستوگرام تفاوت پیکسل‌ها در تصویر استگو دارای توزیعی متفاوت از تصاویر نرمال باشد. تمایز دادن دو هیستوگرام تفاوت تصویر عادی و تصویر استگو شده به یک یا چند پارامتر نیاز دارد. گاهی اوقات با انجام تست‌های مختلف می‌توان حد آستانه مناسبی را برای این پارامترها در تصویر پوشانه و استگو پیدا کرد و با استفاده از این حد آستانه تصاویر پوشانه و استگو را از یکدیگر متمایز کرد.



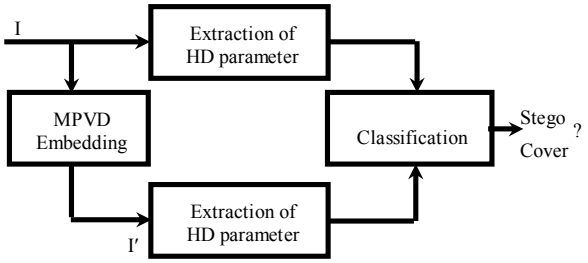
شکل۴: هیستوگرام تفاوت تصویر Lena با توزیع مشابه گوسین

اما گاهی اوقات امکان بدست آوردن این حد آستانه مناسب به آسانی وجود ندارد. راه حلی که در این موارد پیشنهاد می‌شود، استفاده از یک روش یادگیری مانند شبکه عصبی است. در این روش، ابتدا شبکه عصبی با استفاده از پارامتر های تعدادی نمونه آموزشی از هر دو گروه تصاویر استگو و تصاویر پوشانه آموزش داده می‌شود. حال شبکه عصبی آموزش داده شده می‌تواند پارامترهای استخراج شده از یک تصویر تست را دریافت کرده و با دقت مناسب تشخیص دهد که این تصویر یک تصویر پوشانه یا یک تصویر استگو است.

۴- الگوریتم حمله پیشنهادی

با توجه به مباحث انجام شده، گام اول حمله انتخاب پارامترهایی از تصویر است که در اثر جاسازی تغییر محسوسی می‌کند و می‌تواند

متمایز کننده تصویر پوشانه و استگو باشد. با توجه به ویژگی‌های روش MPVD، امیدواریم پارامترهای محاسبه شده از هیستوگرام تفاوت تصویر این تمایز را به خوبی انجام دهند. یک تکنیکی که برای محاسبه این پارامترها در تعدادی از روش‌های حمله موجود استفاده می‌شود، این است که به دنبال پارامترهایی باشیم که تغییر آن‌ها از تصویر عادی به تصویر یک بار استگو شده بسیار باشد، اما با استگو کردن مجدد یک تصویر استگو، این معیارها تغییرات کمی داشته باشند. روند کلی حمله پیشنهادی در شکل ۵ نشان داده شده است.

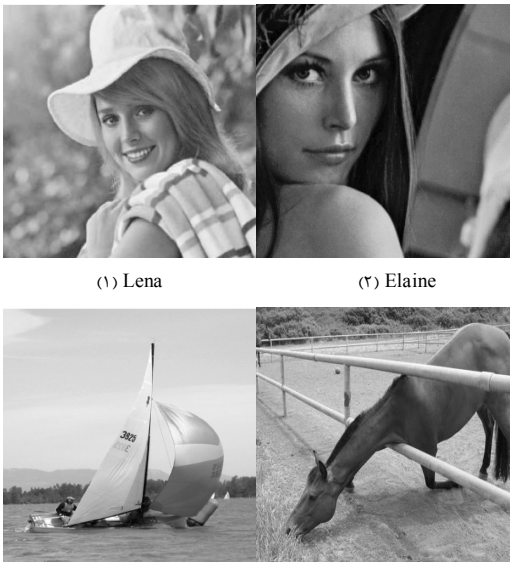


شکل ۵: روند کلی الگوریتم حمله پیشنهادی

فرض کنید می‌خواهیم تصویر مشکوک I را بررسی کنیم و در مورد استگو بودن یا نبودن آن تصمیم گیری کنیم. ابتدا با استفاده از یک عملیات مشخص که در دیاگرام Extraction of HD Parameter نامیده شده است، تعدادی پارامتر مناسب از روی هیستوگرام تفاوت تصویر I محاسبه می‌کنیم. در ادامه این پارامترها را معرفی خواهیم کرد. سپس به روش MPVD در تصویر I جاسازی ۱۰۰٪ انجام می‌دهیم و از روی هیستوگرام تفاوت تصویر تولید شده، I'، پارامترهای موردنظر را محاسبه می‌کنیم. به روشی نیاز داریم که برحسب پارامترهای بدست آمده پوشانه یا استگو بودن تصویر را با دقت خوبی تشخیص دهد.

پس از مشخص شدن الگوریتم کلی حمله پیشنهادی، تنها نکته مبهم تعیین پارامترهای استخراج شده از هیستوگرام تفاوت تصویر است و بررسی اینکه آیا می‌توانیم برای آنها حدآستانه مناسبی بیابیم، یا اینکه این کار را باید به شبکه عصبی واگذار کنیم. برای یافتن پارامترهای مناسب، نیاز است که هیستوگرام‌های تفاوت چند نمونه تصویر تست قبل و بعد از یک و دو بار جاسازی به روش MPVD بررسی شود. به دلیل محدودیت، فقط چند نمونه از تست‌های انجام شده در این جا ارائه می‌شود. در شکل ۶، چند نمونه از تصاویر استاندارد استفاده شده برای محاسبه پارامترهای مورد نظر نشان داده شده است.

در تصاویر تست مختلف دو بار جاسازی کامل به روش MPVD انجام شده است و هیستوگرام های تفاوت تصویر قبل از جاسازی و بعد از هر بار جاسازی در شکل ۷ نشان داده شده است. همانگونه که انتظار داشیم، هیستوگرام تفاوت قبل از جاسازی ، توزیع شبیه گوسین دارد. اگرچه بعد از یک بار جاسازی به روش MPVD، برخلاف PVD ساده، رخداد جفت شدگی اتفاق نیفتاده است، اما توزیع هیستوگرام تفاوت تصویر استگو تولید شده به خوبی حفظ نشده است و تفاوت آشکاری بین دو هیستوگرام قبل و بعد از یک بار جاسازی به روش MPVD

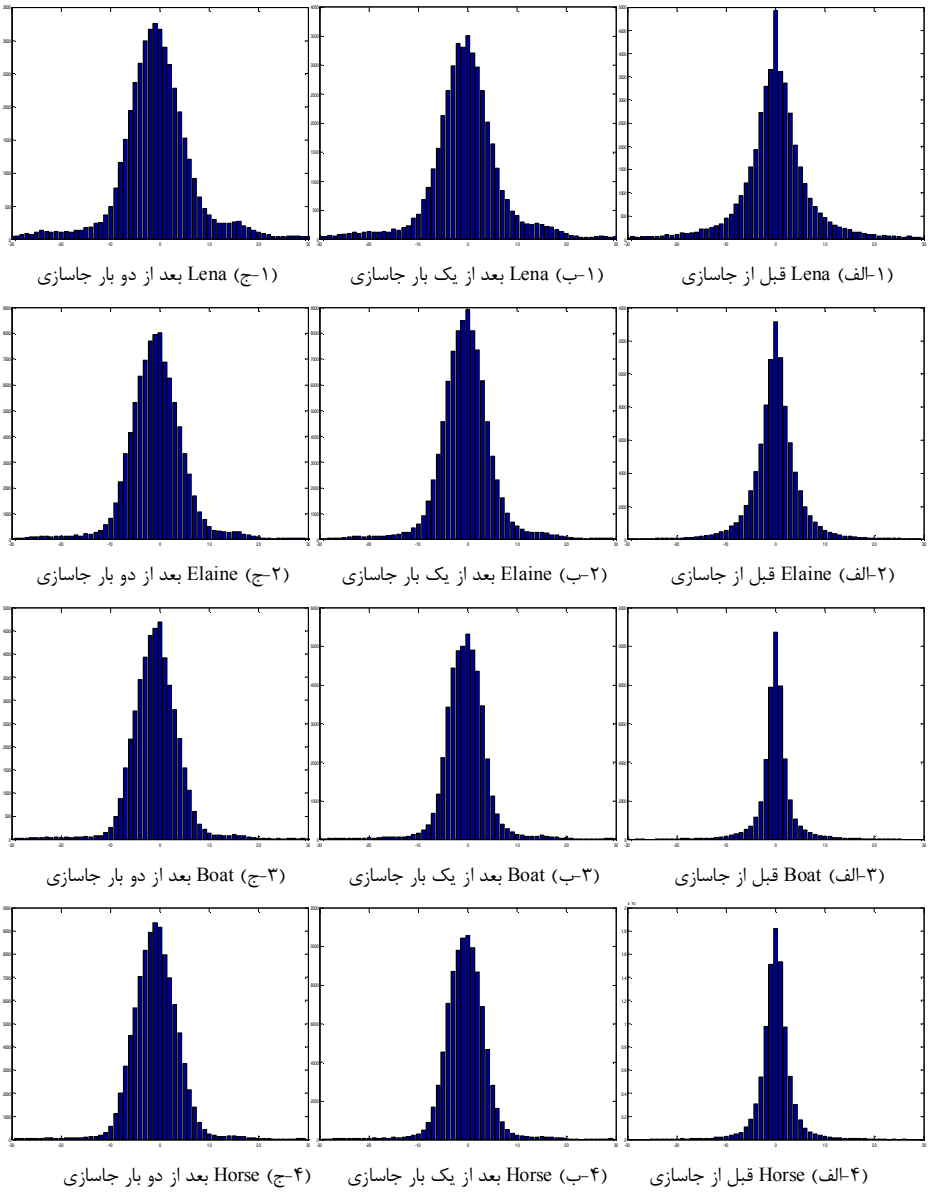

(۱) Lena (۲) Elaine
(۳) Boat (۴) Horse

شکل ۶ : چند نمونه از تصاویر تست

دیده می‌شود. اما هیستوگرام تفاوت‌ها بعد از دو بار جاسازی تقریباً مشابه هیستوگرام تفاوت بعد از یک بار جاسازی است. بنابراین باید بتوانیم پارامترهایی محاسبه کنیم که نشان دهنده تفاوت دو هیستوگرام اول و شباهت دو هیستوگرام دوم باشد.

نکته ای که در مقایسه هیستوگرام‌ها جلب توجه می کند، ارتفاع بلند ستون متناظر با مقدار تفاوت 0 در هیستوگرام های تصاویر عادی و کاهش شدید این ارتفاع در هیستوگرام های تصاویر استگو است. به عبارت دیگر، در هیستوگرام های تصاویر عادی، تفاوت ستون 0 با دو ستون مجاور ۱- و ۱ بسیار بیشتر از این تفاوت در هیستوگرام های تفاوت تصاویر یک بار و دو بار استگو شده است. همین اتفاق برای ستون‌های مجاور (۱ و ۲) و (۱- و ۲-) نیز رخ می‌دهد. برای مثال در تصویر عادی Boat، مقدار تفاوت های۲-، ۱- ، ۰، ۱ و ۲ به ترتیب ۴۱۴۳، ۷۸۸۳، ۱۰۷۵۰، ۷۹۴۶ و ۴۱۶۸ بوده است. اما بعد از یک بار جاسازی تعداد این تفاوت ها تبدیل به ۴۸۷۹، ۵۰۰۳، ۵۳۱۲، ۴۹۰۰ و ۴۳۵۲ شده‌اند. بدین ترتیب کاهش ستون ۰ از ۱۰۷۵۰ به ۵۳۱۲ و نزدیک شدن مقادیر ۱- و ۱ به ستون ۰ و مقادیر ۲ و ۲- به ۱و ۱- در تصویر یک بار استگو شده به خوبی می‌تواند دو تصویر را از هم مجزا کند. اما بعد از دو بار جاسازی، مقادیر این ستون‌ها برابر ۴۳۸۵، ۴۵۸۵، ۴۶۵۷، ۴۰۱۳ و ۳۴۳۰ شده اند. اگرچه بازهم کاهش مقادیر ستون‌ها رخ داده است، اما این کاهش در مقایسه با کاهش مرحله قبل شدید نیست

و تفاوت بین ستون‌ها نیز تقریباً همانند مرحله قبل است.



شکل ۷: هیستوگرام تفاوت چند تصویر تست نمونه

بنابراین با توجه به مقایسه‌های انجام شده، به نظر می‌رسد که می‌توان پارامتری براساس مقادیر سه ستون ۰، ۱ و ۱- در هیستوگرام تفاوت تصاویر محاسبه کرد و با انجام تست‌های گوناگون حد آستانه‌ای برای این پارامتر پیدا کرد که بتواند میان تصاویر عادی و استگو تمایز ایجاد کند. این پارامتر را می‌توان به روش‌های مختلف تعریف کرد، اما در این بخش تنها یکی از این فرمول‌ها بررسی می‌شود. فرض کنید d_0 ، d_1 و d_{-1} به ترتیب مقادیر ستون‌های ۰، ۱ و ۱- در هیستوگرام تفاوت تصویر باشند، پارامتر P به صورت زیر قابل محاسبه است:

$$P = \frac{d_0 - (\frac{d_1 + d_{-1}}{2})}{d_0 + d_1 + d_{-1}} \quad (۷)$$

در پارامتر P، تفاوت ستون 0 با میانگین ستون‌های ۱ و ۱- مورد توجه قرار گرفته است و برای جلوگیری از بزرگ شدن این پارامتر، مقدار موردنظر به مجموع سه ستون نرمالیزه شده است. بدین ترتیب انتظار می‌رود که برای تصاویر عادی که تفاوت ستون ۰ با ستون‌های کناری زیاد است، مقدار این پارامتر بزرگ باشد و برای تصاویر استگو مقدار این

پارامتر کوچک باشد.

برای تصمیم گیری در مورد کوچک یا بزرگ بودن این پارامتر، یک بار برای تصویر مشکوک مقدار پارامتر P محاسبه شده و PC نامگذاری می‌شود. سپس در تصویر به روش MPVD جاسازی می‌شود و مجدداً پارامتر P محاسبه شده برای آن PS نامیده می‌شود. اگر تصویر مشکوک یک تصویر عادی باشد، PC بزرگ و PS کوچک است. پس نسبت آنها یک عدد بزرگ می‌شود. اما اگر تصویر از ابتدا استگو بوده باشد، PC و PS کوچک و تقریباً مساوی می‌شوند، پس انتظار می‌رود که این نسبت یک عدد کوچک و نزدیک به ۱ باشد.

در جدول ۱، لیستی از پارامترهای استفاده شده برای آموزش شبکه عصبی ارائه شده است. اگر تصویر I، تصویر مورد نظر برای تشخیص استگو یا پوشانه بودن، و تصویر I'، تصویر حاصل از انجام فرآیند جاسازی به روش MPVD در I باشد. d_i و d'_i، به ترتیب ستون i ام هیستوگرام تفاوت تصویر I و I' را نشان می‌دهد. همانگونه که در شکل‌ها دیدیم، ستون‌های ۲- تا ۲ فقط برای این هدف کافی است و می‌توانیم از مابقی ستون‌ها صرفنظر کنیم. علاوه بر مقدار این ستون‌ها، مقدار تفاوت ستون‌های مجاور نیز در هر دو تصویر به عنوان پارامتر انتخاب شده‌اند که پارامترهای گروه ۲ و ۴ آنها را نشان می‌دهند. بعلاوه دو پارامتر PC و PS و نسبت این دو پارامتر نیز انتخاب شده‌اند.

جدول ۱: پارامترهای استخراج شده از تصویر

تعداد	پارامتر	گروه
۵	$d_i \ (-2 \le i \le 2)$	۱
۴	$d_{i+1} - d_i \ (-2 \le i \le 1)$	۲
۵	$d'_i \ (-2 \le i \le 2)$	۳
۴	$d'_{i+1} - d'_i \ (-2 \le i \le 1)$	۴
۱	PC	۵
۱	PS	۶
۱	PC/PS	۷

بدین ترتیب گام اول الگوریتم، انتخاب پارامترها پایان می‌یابد. در گام دوم باید از یک رده بند مناسب برای تعیین نوع تصویر ورودی استفاده کرد. با توجه به تست‌های انجام شده، برای این کار از یک شبکه عصبی MLP [9] استفاده شده است. مشخصات این شبکه در جدول ۲ آورده شده است. در بخش بعدی نتایج تست‌های انجام شده و دقت الگوریتم موردنظر ارائه می‌شود.

[9] Multi Layer Perceptron (MLP)

۵- نتایج پیاده سازی

پیاده‌سازی الگوریتم پیشنهادی در دو مرحله انجام شده است. گام اول، استخراج پارامترها، در Matlab 7.1 و گام دوم، شبکه عصبی، با کمک نرم افزار Neurosolution 5 پیاده‌سازی شده است. برای آموزش شبکه عصبی به مجموعه نسبتاً بزرگی از پارامترهای استخراج شده از تصاویر هر دو گروه نوع پوشانه و استگو نیاز است. برای جمع آوری این مجموعه داده آموزشی و تست، از ۳۵۰ تصویر پوشانه و استگو استفاده شده است. ۸۰٪ از این مجموعه برای آموزش و ۲۰٪ باقیمانده برای تست انتخاب

جدول ۲: مشخصات شبکه عصبی پیشنهادی

ویژگی	مقدار
نوع شبکه	MLP
الگوریتم یادگیری	Back Propagation
تعداد ورودی	۲۱
تعداد لایه های مخفی	۱
تعداد نرون لایه مخفی	۸
تابع انتقال نرون لایه مخفی	Hyperbolic tangent
تعداد نرون لایه خروجی	۱
تابع انتقال نرون لایه خروجی	Linear Hyperbolic tangent

شده است. بنابراین ابتدا از ۱۴۰ تصویر استگو و ۱۴۰ تصویر پوشانه پارامترهای لیست شده در جدول ۱ استخراج شد و شبکه عصبی‌های مختلف با تعداد پارامترهای ورودی مختلف (از مجموعه پارامترهای لیست شده) آموزش داده شد. سپس پارامترهای استخراج شده ۷۰ تصویر تست باقیمانده که ۳۵ تصویر پوشانه و ۳۵ تصویر استگو است، به ورودی شبکه عصبی آموزش داده شده، داده شد. نتایج خروجی این شبکه برای مجموعه تست نشان دهنده دقت الگوریتم است.

برای اینکه شبکه عصبی با مشخصات مناسب و مجموعه پارامترهای ورودی مناسب مشخص شود، تست‌های مختلفی انجام شده است. در واقع برای شبکه عصبی‌های با مشخصات متفاوت و با مجموعه ورودی-های متفاوت دقت رده بندی محاسبه شده است. این نتایج در جدول ۳ نشان داده شده است. TP[10]، تعداد تصاویر استگویی است که به درستی کشف شده است و FP[11]، تعداد تصاویر پوشانه‌ای است که به اشتباه استگو اعلام شده است. بنابراین هر چه TP بزرگتر و FP کوچکتر باشد، شبکه مناسب‌تر است.

[10] True Positive (TP)

[11] False Positive (FP)

نتایج ارائه شده نشان می‌دهد که بهترین حالت زمانی است که تمام پارامترهای جدول ۱ به عنوان ورودی به شبکه عصبی MLP داده شود و شبکه دارای یک لایه مخفی با ۸ نرون و با مشخصات لیست شده در جدول ۲ باشد. تست‌های مشابهی نیز برای مقایسه توابع انتقالی مختلف و انواع شبکه‌ها انجام شده است که مجالی برای ارائه تمام نتایج تست در اینجا وجود ندارد. اگر P و N به ترتیب تعداد تصاویر استگو و پوشانه موجود در مجموعه تست باشد، دقت[12]رده بندی را به صورت زیر می-توان محاسبه کرد :

$$Accuracy=\frac{(TP+TN)}{(P+N)} \qquad (۸)$$

TN، تعداد تصاویر پوشانه‌ای است که به درستی تشخیص داده شده است. بنابراین با توجه به تعریف دقت و نتایج ارائه شده در جدول ۳، دقت شبکه پیشنهادی ۹۸/۶٪ است.

جدول ۳: دقت الگوریتم با شبکه های با ساختارهای مختلف

شبکه / گروه ورودی	تعداد ورودی	لایه مخفی با ۴ نرون		لایه مخفی با ۸ نرون	
		TP	FP	TP	FP
۱	۵	۲۹	۱	۲۹	۱
۱و۲	۹	۳۰	۵	۳۱	۴
۱و۳	۱۰	۳۱	۰	۳۲	۰
۱و۲و۳و۴	۱۸	۳۴	۱	۳۵	۳
۱و۲و۳و۴و۵و۶	۲۰	۳۵	۲	۳۵	۲
تمام گروه ها	۲۱	۳۵	۲	۳۵	۱

۶- نتیجه گیری

روش MPVD[۱۳]، بر خلاف روش‌های LSB-F و LSB-M، برای جاسازی از بیت‌های کم ارزش پیکسل‌های تصویر به صورت مستقیم استفاده نمی‌کند و جاسازی را در مقدار تفاوت پیکسل‌های درون بلاک-های دوتایی انجام می‌دهد. به همین دلیل حملاتی که با بررسی هیستوگرام تصویر تلاش می‌کنند وجود پنهان‌نگاری را اثبات کنند، قادر به کشف این روش نیستند.

با توجه به این‌که جاسازی در مقدار تفاوت‌ها انجام می‌شود، تصاویر استگو دارای توزیعی متفاوت از تصاویر نرمال هستند. برای تشخیص و تمایز دو هیستوگرام تفاوت تصویر عادی و تصویر استگو شده مجموعه-ای از ۲۱ پارامتر پیشنهاد شده است که تمام پارامترها از هیستوگرام تفاوت تصویر استخراج شده است. بعد از انجام تست‌های مختلف، که در بخش‌های قبل به تعدادی از آنها اشاره شد، یک شبکه عصبی MLP با یک لایه مخفی با ۸ نرون برای انجام رده‌بندی پیشنهاد شد. این شبکه با پارامترهای استخراج شده از ۲۸۰ تصویر استگو و پوشانه آموزش داده شده است. شبکه آموزش داده شده، توانسته است مجموعه تستی شامل ۷۰ تصویر استگو و پوشانه را با دقت ۹۸/۶٪ رده بندی کند.

شبکه پیشنهاد شده قادر به رده بندی تصاویر استگویی است که به صورت کامل به روش MPVD در آن داده جاسازی شده است. برای ادامه این کار می‌توان در مورد ساختار شبکه و پارامترهای ورودی جستجو کرد که قادر به رده بندی مناسب تصاویری باشد که کمتر از ۱۰۰٪ در آنها داده جاسازی شده است.

مراجع

[1] Wayner, p., "Disappearing cryptography", *Elsivier Science*, 2002.

[2] Anderson, R.J., Petitcolas F.A.P., "On the limits of steganography", *IEEE Journal of Selected Areas in Communications, Special Issue on Copyright and privacy Protection*, Vol. 16, No. 4, pp. 474-481, 1998.

[3] Pitzmann, B., "Information hiding terminology", *in Information Hiding, Springer Lecture Notes in Computer Science*, Vol. 1174, pp. 347-350, 1996.

[4] Provos, N., Honeyman,P., "Hide and seek: an introduction to steganography", *IEEE Security & Privacy*, pp. 32-44, 2003.

[5] Kharrazi, M., Sencar, H.T., Memon, N., "Image steganography: concepts and practice", *WSPC/ Lecture Notes Series*, 2004.

[6] Ker, A.D., "Steganalysis of LSB matching in grayscale images", *IEEE Signal Processing Letters*, Vol. 12, No. 6, 2005.

[7] Westfeld, A., Pfitzman, A., "Attacks on steganographic systems", *Proc. 3rd Int'l Information Hiding Workshop*, pp. 61-76, 1999.

[8] Fridrich, J., Goljan, M., Du, R., "Reliable detection of LSB steganography in grayscale and color images", *Proc. of ACM: Special Session on Multimedia Security and Watermarking*, 2001.

[9] Ker, A., "Resampling and the detection of LSB matching in color bitmaps", *Security, Steganography and Watermarking of Multimedia Contents VII, Proceedings of SPIE 5681*, pp. 1-15, 2005.

[10] Wu, D.C, Tsai, W.H., "A steganographic method for images by pixel-value differencing", *Pattern Recognition Letters*, vol. 24, 2003.

[11] Wu, H.C, Wu, N.I, Tsai, C.S, Hwang, M.S., "Image steganographic scheme based on pixel-value differencing and LSB replacement methods", *IEEE Proceedings Vision, Image and Signal Processing*, Vol. 152, No. 5, pp. 611-615, 2005.

[12]Sabeti, V., Samavi, S., Mahdavi, M., Shirani, S., "Steganalysis of pixel-value differencing steganographic method", *IEEE Pacific Rim Conference on Communications, Computers and Signal Processing*, pp. 292-295, 2007.

[12] Accuracy

[13]Zhang, Z., Wang, S., "Vulnerability of pixel-value differencing steganography to histogram analysis and modification for enhanced security", *Pattern Recognition Letters*, No. 3, pp. 331-339, 2004.

[14]Farid, H., "Detecting Steganographic Message in Digital Images", *Report TR2001-412*, Dartmouth College, Hanover, NH, 2001.

[15]Fridrich, J., Pevny, T., "Merging Markov and DCT features for multi-class JPEG steganalysis", *Proc. SPIE Electronic Imaging, Photonics West*, pp. 03-04, 2007.

[16]Liu, Q., Sung, A.H., Chen, Z., Xu, J., "Feature mining and pattern classification for steganalysis of LSB matching steganography in grayscale images", *Pattern Recognition*, Vol. 41, pp. 56 – 66, 2008.

[17]Zhang, T., Ping, X., "A new approach to reliable detection of LSB steganography in natural images", *Signal Processing 83*, 2003.