

# آشکارسازی جعل و فشرده‌سازی مجدد در تصاویر دیجیتال با استفاده از ترکیب ویژگی‌های مبتنی بر اثر کوانتیزاسیون بر ضرایب DCT

ندا علی‌پور<sup>۱</sup>، دانشجوی کارشناسی ارشد؛ علیرضا بهراد<sup>۲</sup>، دانشیار

۱- دانشکده فنی مهندسی - دانشگاه شاهد - تهران - ایران - nedalipur@gmail.com

۲- دانشکده فنی مهندسی - دانشگاه شاهد - تهران - ایران - behrad@shahed.ac.ir

**چکیده:** آشکارسازی فشرده‌سازی مجدد، نقش مهمی در جرم‌کاوی تصاویر JPEG و تشخیص جعل دارد. در این مقاله، روش جدیدی برای تشخیص جعل بر اساس آشکارسازی فشرده‌سازی مجدد تصاویر دیجیتال با استفاده از ترکیب ویژگی‌های مبتنی بر اثر کوانتیزاسیون بر ضرایب DCT ارائه می‌شود. در روش پیشنهادی از ترکیب ویژگی‌های مبتنی بر توزیع ارقام اول ضرایب DCT و همچنین ویژگی‌های مبتنی بر پیوند کوانتیزاسیون در ضرایب DCT هر سه کانال رنگی تصویر استفاده می‌شود. به منظور انتخاب ویژگی‌های مناسب و نیز کاهش حجم محاسبات ما از یک الگوریتم سلسله مراتبی انتخاب ویژگی برای کاهش تعداد ویژگی‌ها و انتخاب ویژگی‌های مناسب استفاده می‌کنیم. این روش علاوه بر انتخاب ناحیه فرکانسی مناسب، امکان انتخاب توزیع رقم با بیشترین تأثیر را نیز فراهم می‌کند. برای آزمایش الگوریتم پیشنهادی و مقایسه آن با سایر روش‌ها از پایگاه داده تصاویر جعلی و غیرجعلی CASIA استفاده می‌شود. نتایج آزمایش‌ها و مقایسه روش پیشنهادی با سایر روش‌های موجود، نشانگر صحت شناسایی ۹۹/۵۵ درصد و بهبود ۷/۶ درصد نسبت به روش‌های موجود است که بیانگر عملکرد مناسب الگوریتم پیشنهادی است.

**واژه‌های کلیدی:** جرم‌کاوی تصویر دیجیتال، فشرده‌سازی مجدد، تشخیص جعل، تصویر JPEG

## Forgery and Double Compression Detection in Digital Images using Combined Features of Quantization Effects on DCT Coefficients

N. Alipour<sup>1</sup>, MSc Student; A. Behrad<sup>2</sup>, Associate Professor

1- Faculty of Engineering, Shahed University, Tehran, Iran, Email: nedalipur@gmail.com

2- Faculty of Engineering, Shahed University, Tehran, Iran, Email: behrad@shahed.ac.ir

**Abstract:** Recompression detection has an important role in JPEG image forensic and forgery detection. In this paper, a new algorithm is proposed for image forgery detection using combined features of quantization effect on DCT coefficients. In the proposed approach, we use distribution of most significant digit of DCT coefficients and features based on the period of DCT coefficients in three color channels. To select more proper features and reduce computational overhead of the algorithm, a new hierarchical approach is used for feature selection. The proposed feature selection algorithm can select both proper frequency region and more effective digit distribution for efficient forgery detection. To test the proposed algorithm and compare the results with those of other methods the forged and authentic images of CASIA database are used. The experimental results and a comparison of the results of the proposed algorithm with those of other method show the correct recognition rate of 99.5% as well as the enhancement of 7.6% with respect to existing approaches which demonstrate the efficiency of the proposed approach.

**Keywords:** Digital image forensic, double compression, forgery detection, JPEG image.

تاریخ ارسال مقاله: ۱۳۹۴/۱۰/۱۵

تاریخ اصلاح مقاله: ۱۳۹۴/۱۲/۲۶

تاریخ پذیرش مقاله: ۱۳۹۵/۰۳/۲۰

نام نویسنده مسئول: علیرضا بهراد

نشانی نویسنده مسئول: ایران - تهران - بزرگراه خلیج فارس - دانشگاه شاهد - دانشکده فنی مهندسی.

## ۱- مقدمه

کرد. این روش‌ها معمولاً برای تصاویر با فشرده‌سازی کم خوب عمل می‌کنند اما برای تصاویر با فشرده‌سازی بالا نتایج معمولاً مناسب نیستند.

از اثر هرگونه عملیاتی که برای ویرایش تصویر استفاده می‌شود نیز می‌توان به جعل تصویر پی برد. البته برخی عملیات همانند فیلتر کردن و تیزسازی تصویر لزوماً به معنای جعل نیستند ولی عملیاتی مانند کپی کردن، همانندسازی و انتقال قسمتی از تصویر به همان تصویر یا تصویر دیگر که به پیوندزنی<sup>۲</sup> تصویر معروف است، به‌منزله جعل هستند. تشخیص جعل با استفاده از تشخیص همانندسازی و پیوندزنی در مراجع متعددی مورد استفاده قرار گرفته است [۸-۱۳].

تصویربرداری نگاشت فضای سه‌بعدی به فضای دوبعدی با استفاده از مدل افکنشی پرسپکتیو و قوانین اپتیکی عدسی است. چشم انسان معمولاً به تغییرات هندسی زیاد حساس نیست، لذا ناسازگاری‌های هندسی ناشی از جعل به‌راحتی توسط چشم انسان آشکارسازی نمی‌شوند. یکی از روش‌های تشخیص جعل تصاویر دیجیتال استفاده از تشخیص ناسازگاری‌های فیزیکی، هندسی یا محتوای تصویر است. برای نمونه‌ای از این روش‌ها می‌توان به روش تشخیص ناسازگاری‌های سایه [۱۴]، ناسازگاری‌ها هندسی و نور تصویر [۱۵، ۱۶] و ناسازگاری‌های حرکتی در ویدئو [۱۷] اشاره کرد. روش تشخیص ناسازگاری‌های فیزیکی، هندسی یا محتوای تصویر برای تصاویر با محتوای خاص قابل‌اعمال هستند و عمومیت ندارند.

امروزه فشرده‌سازی یکی از مراحل اجتناب‌ناپذیر نرم‌افزاری پردازش و ویرایش تصویر هستند. فشرده‌سازی اثرات خاص به خود را در تصویر به‌جای می‌گذارد لذا استفاده از اثرات فشرده‌سازی و کدینگ تصویر یکی از روش‌های تشخیص جعل محسوب می‌شود.

قالب JPEG<sup>۱</sup> یکی از پرکاربردترین قالب‌های مورد استفاده در کدینگ و فشرده‌سازی تصویر است. بنابراین آشکارسازی جعل در تصاویر JPEG، نقش مهمی را در تشخیص دست‌کاری‌های تصویری ایفا می‌کند. از آنجایی که بیشتر تصاویر JPEG دست‌کاری شده، متحمل حداقل دو بار فشرده‌سازی می‌شوند، بنابراین یکی از روش‌های بسیار مهم اعتبارسنجی تصویر JPEG، مبتنی بر آشکارسازی فشرده‌سازی مجدد است.

در [۱۸] از نویز کوانتیزاسیون برای بررسی فشرده‌سازی تصاویر JPEG استفاده شده است. در مرجع [۱۹] نیز تأثیر فشرده‌سازی مضاعف بر هیستوگرام ضرایب DCT بررسی شده است. در این مقاله نشان داده شده است که فشرده‌سازی مضاعف با ضرایب متفاوت باعث ایجاد الگوهای پررنگ در هیستوگرام می‌شود که می‌تواند در حوزه فوریه آشکارسازی شود. از اشکالات این روش تشخیص فشرده‌سازی با استفاده از کاربر انسانی است. در مراجع [۲۰، ۲۱] ضمن بسط روش الگوهای پررنگ در هیستوگرام از ویژگی‌های استخراج شده و طبقه‌بند ماشین بردار پشتیبان (SVM) برای تشخیص خودکار و بدون کاربر انسانی استفاده شده است.

امروزه پردازش تصویر در زمینه‌های مختلفی از قبیل کاربردهای صنعتی، ورزشی [۱] و پزشکی [۲] استفاده می‌شود. اما با رشد سریع ابزارهای پردازش تصویر جعل تصویر نیز گسترش یافته است. جعل تصویر تاریخچه طولانی دارد اما در عصر حاضر که عصر دیجیتال است، با رشد سریع ابزارهای پردازش تصویر، دست‌کاری یک تصویر دیجیتال بدون به‌جا گذاشتن اثری که قابل‌مشاهده باشد به‌آسانی انجام می‌شود. لذا تغییر اطلاعات و محتوای تصویر و ایجاد تصویری که به‌ظاهر معتبر باشد، به‌راحتی صورت می‌گیرد. بررسی صحت تصاویر دیجیتال و همچنین آشکارسازی اثرات و ردپاهای دست‌کاری و جعل تصاویر زمینه مهم تحقیقاتی در بحث پردازش تصویر محسوب می‌شود. تعیین صحت تصاویر و عدم جعل آن‌ها، نقشی مهم در بسیاری از زمینه‌ها داشته که شامل مواردی همچون سیستم‌های نظارت، جرم‌کاوی، خدمات هوشمند، تصویربرداری پزشکی و روزنامه‌نگاری است.

تحول و دگرگونی اطلاعات دیجیتالی و همچنین موضوعاتی که مرتبط با امنیت رسانه‌ای و امنیت اطلاعات می‌باشند، منجر به ایجاد روش‌هایی در آشکارسازی جعل شده است، که این روش‌ها به‌صورت کلی به دو دسته روش‌های فعال و غیرفعال دسته‌بندی می‌شود. آبنشانی<sup>۱</sup> دیجیتالی نمونه‌ای از روش‌های فعال است. این روش‌ها که نیاز به درج کدهای اولیه یا آبنشان در تصویر دارند، به‌طور عام قابل‌استفاده نیستند و تجهیزات منحصربه‌فردی لازم دارند لذا استفاده عمومی آن‌ها برای تشخیص جعل در تصاویر بدون آبنشان امکان‌پذیر نیست.

بیشتر روش‌های نوین جرم‌کاوی تصویر از آشکارسازی سرخ‌های معتبری در حین ایجاد و دست‌کاری تصویر بهره می‌برند که به روش‌های غیرفعال مشهورند. در مقایسه با روش‌های فعال، روش‌های غیرفعال نیازی به آبنشان یا امضای اضافی ندارند و می‌توانند کاربردهای گسترده‌تر و عمومی‌تری در تشخیص جعل تصویر داشته باشند. روش‌های غیرفعال بر اساس ویژگی‌های استفاده‌شده برای جرم‌کاوی به گروه‌های مختلفی گونه‌بندی می‌شوند که از جمله می‌توان به ۱- روش‌های مبتنی بر سرخ‌های وسیله تصویربرداری اخذ تصویر، ۲- روش‌های مبتنی برای اثرهای ویرایشی، ۳- روش‌های مبتنی بر ناسازگاری‌های فیزیکی، هندسی یا محتوا و ۴- روش‌های مبتنی بر سرخ‌های کدینگ تصویر اشاره کرد.

به‌منظور اخذ تصویر و ذخیره‌سازی آن توسط دوربین مراحل متعددی توسط دوربین صورت می‌گیرد که این مراحل و نوع آن ممکن است از یک دوربین به دوربین دیگری فرق کند. در برخی از الگوریتم‌های تشخیص جعل از اثرات مراحل تصویربرداری بر تصویر خروجی استفاده می‌شود. برای نمونه‌ای از این روش‌ها می‌توان به روش‌های مبتنی بر نویز غیریکنواخت پاسخ تصویر (PRNU)<sup>۲</sup> [۳]، روش انحراف رنگی [۴، ۵] و نیز روش میان‌بایی رنگی [۶، ۷] اشاره

یکی دیگر از روش‌های تشخیص فشرده‌سازی مضاعف در تصاویر JPEG استفاده از روش توزیع آماری ارقام اول ضرایب تبدیل کسینوسی گسسته (DCT<sup>s</sup>) با استفاده از قانون بنفورد است [۲۴-۲۲]. قانون بنفورد بیان می‌کند که توزیع احتمال اولین ارقام  $x$  در یک مجموعه از ارقام طبیعی، لگاریتمی هستند. وقتی که تصویر یک‌بار فشرده شده باشد توزیع مهم‌ترین ارقام ضرایب بلوک DCT از قانون بنفورد پیروی می‌کنند ولی این توزیع، به دو بار فشرده‌گی، بسیار حساس است، به عبارتی قانون لگاریتمی بعد از دو بار فشرده‌گی دچار اغتشاش می‌شود که از این قانون برای تشخیص فشرده‌سازی مجدد استفاده می‌شود.

## ۲-۱- استخراج ویژگی

ما در این مقاله، روش جدیدی برای آشکارسازی جعل و فشرده‌سازی مضاعف با استفاده از تلفیق ویژگی‌های به‌دست‌آمده از روش‌های توزیع ارقام بارزش ضرایب DCT و همچنین ویژگی‌های مبتنی بر پیوند کوانتیزاسیون ضرایب DCT استفاده کرده‌ایم. ما هم از اطلاعات رنگی و هم اطلاعات شدت روشنایی برای استخراج ویژگی استفاده کرده ولی به‌منظور کاهش حجم محاسبات و همچنین افزایش صحت الگوریتم پیشنهادی از یک الگوریتم سلسله مراتبی برای انتخاب ویژگی‌های مناسب استفاده کردیم.

ساختار این مقاله نیز به شرح زیر است: در بخش دوم مراحل مختلف الگوریتم پیشنهادی ارائه و شرح داده می‌شود. بخش سوم به نتایج آزمایشی و مقایسه الگوریتم با روش‌های قبلی می‌پردازد و در نهایت در بخش چهارم نیز نتیجه‌گیری مقاله آورده می‌شود.

### ۲-۱-۱- ویژگی‌های مبتنی بر پیوند هیستوگرام ضرایب DCT

شکل ۲ هیستوگرام ضرایب DCT را به ازای یک بار فشرده‌سازی و دو بار فشرده‌سازی با ضرایب کوانتیزاسیون مختلف را نشان می‌دهد. همان‌طور که در شکل نشان داده شده است در صورت فشرده‌سازی مجدد ضرایب DCT هیستوگرام این ضرایب دارای الگوهای پیوندیک با پیوندی غیر از یک خواهد بود. در مرجع [۱۹] از تحلیل حوزه فوریه توسط کاربر انسانی برای تشخیص پیوند و فشرده‌سازی مضاعف استفاده شده است. این روش خودکار نبوده و نیازمند تحلیل انسانی می‌باشد.

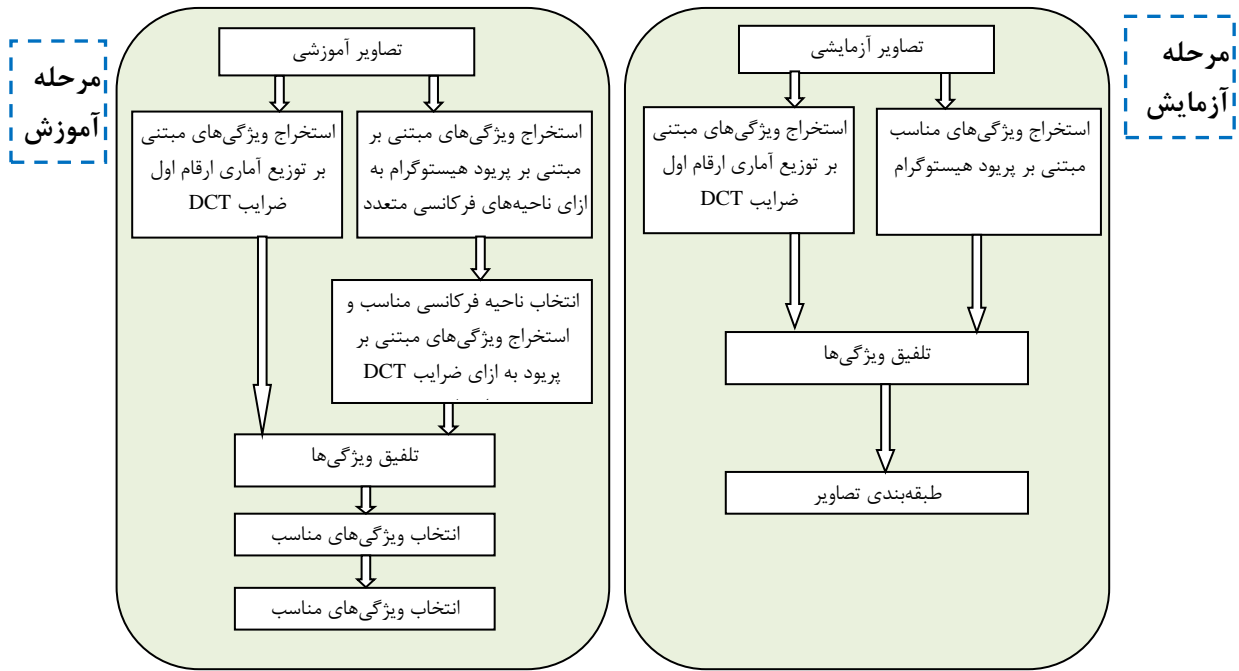
در الگوریتم پیشنهادی از ویژگی‌های مرجع [۲۰] استفاده شده است ولی برای افزایش صحت الگوریتم، ضرایب DCT و ناحیه فرکانسی به‌صورت پویا و سازگار انتخاب می‌شوند. ویژگی‌های این مرحله از الگوریتم، مبتنی بر محاسبه پیوند هیستوگرام است. ما برای محاسبه پیوند هیستوگرام از دو روش استفاده‌شده در مراجع [۲۰] و [۲۱] بهره بردیم. در مرجع [۲۰] پیوند با بیشینه کردن رابطه زیر محاسبه می‌شود.

$$H(p) = \frac{1}{i_{\max} - i_{\min} + 1} \sum_{i_{\min}}^{i_{\max}} h(i \times p + s_0) \quad (1)$$

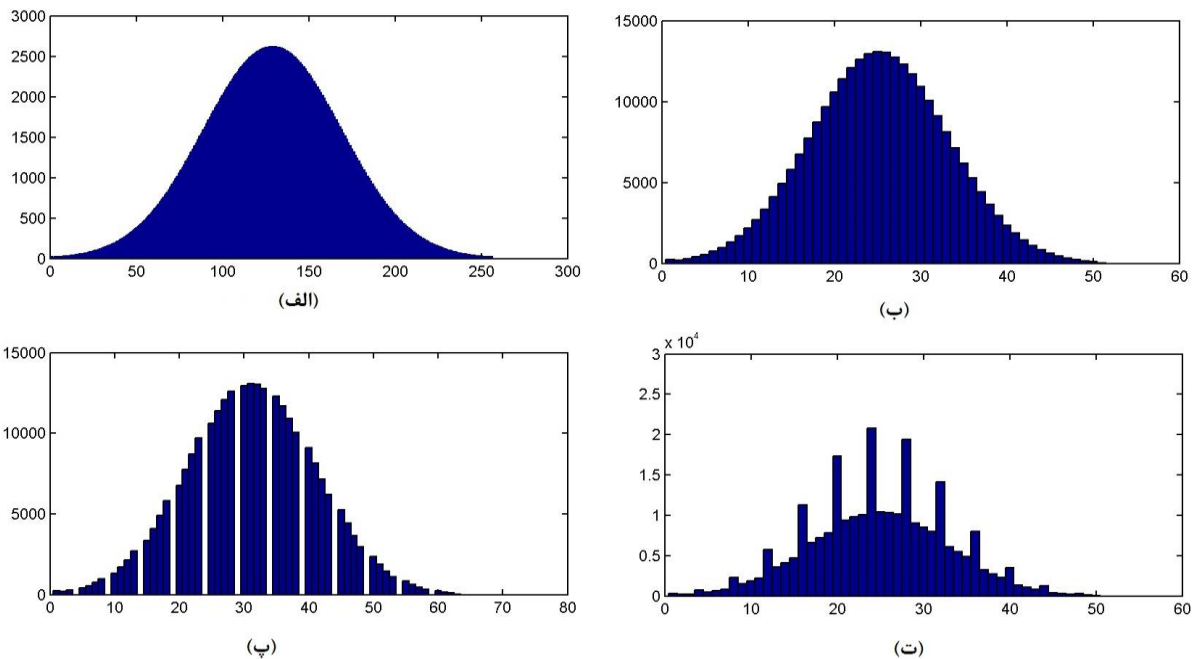
در رابطه بالا  $h$  نمایشگر هیستوگرام ضریب DCT،  $p$  پیوند هیستوگرام،  $s_0$  اندیس بازه‌ای از هیستوگرام با بیشترین مقدار است.  $i_{\min}$  و  $i_{\max}$  نیز از روابط زیر محاسبه می‌شوند [۲۰]:

$$i_{\max} = \frac{s_{\max} - s_0}{p} \quad (2)$$

شکل ۱ بلوک دیاگرام کلی الگوریتم پیشنهادی را نشان می‌دهد، در این مقاله، روش جدیدی برای آشکارسازی جعل با استفاده از تشخیص فشرده‌سازی مجدد در تصاویر JPEG ارائه می‌شود. همان‌طور که در شکل ۱ نشان داده شده است الگوریتم از دو قسمت آموزش و آزمایش تشکیل شده است. قسمت آموزش به‌صورت برون‌خط صورت گرفته و شامل آموزش طبقه‌بند و استخراج ویژگی‌های مناسب است. در روش‌های پیشنهادی از تلفیق دو نوع ویژگی شامل ویژگی‌های مبتنی بر پیوند هیستوگرام ضرایب DCT و نیز توزیع آماری ارقام اول ضرایب DCT استفاده شده است. یکی از عوامل مهم در عملکرد مناسب این دو ویژگی، انتخاب مناسب ضرایب DCT و ناحیه مناسب فرکانسی برای کانال‌های مختلف رنگی است. در الگوریتم پیشنهادی، انتخاب ناحیه مناسب فرکانسی برای استخراج ویژگی مناسب در دو مرحله صورت می‌گیرد. در مرحله اول ناحیه مناسب فرکانسی برای تعیین ویژگی‌های مبتنی بر پیوند هیستوگرام ضرایب DCT صورت می‌گیرد تا ضرایب مناسب DCT برای تعیین این نوع ویژگی‌ها استخراج شوند. در مرحله بعدی توزیع آماری ارقام اول ضرایب DCT برای کلیه ضرایب DCT کانال‌های رنگی محاسبه شده و ویژگی‌های حاصل از آن با ویژگی‌های مبتنی بر پیوند



شکل ۱: بلوک دیاگرام کلی الگوریتم پیشنهادی



شکل ۲: تأثیر فشرده‌سازی مجدد بر پریود هیستوگرام ضرایب DCT: (الف) هیستوگرام اولیه بدون فشرده‌سازی، (ب) هیستوگرام با یک بار کوانتیزاسیون با ضریب ۵، (پ) هیستوگرام با دو بار کوانتیزاسیون به ترتیب با ضرایب ۵ و ۴، (ت) هیستوگرام با دو بار کوانتیزاسیون به ترتیب با ضرایب ۵ و ۴

$$f(i, p) = \frac{h(i) - \text{Max}\{N(z, p)\}}{\text{Mean}\{N(z, p)\}} \quad (5)$$

$$i_{\min} = \frac{s_{\min} - s_0}{p} \quad (3)$$

که  $s_{\min}$  و  $s_{\max}$  به ترتیب کمترین و بیشترین اندیس در هیستوگرام ضریب DCT هستند. در مرجع [۲۱] نیز پریود با بیشینه کردن  $H(p)$  محاسبه می‌شود که روابط محاسبه  $H(p)$  به شرح زیر هستند:

$$H(p) = \text{Median}\{f(i, p) | i = x \times p + s_0, i_{\min} \leq x \leq i_{\max}\} \quad (6)$$

$$N(z, p) = \{h(i) | z - p < i < z + p, i \neq z\} \quad (4)$$

$$P(\text{tampered} | s_0 + i) = \frac{P_t}{(P_t + P_u)} \quad (9)$$

$$P(\text{unchanged} | s_0 + i) = \frac{P_t}{(P_t + P_u)} \quad (10)$$

با جمع احتمال‌های پسین برای همه پیکسل‌های یک بلوک از تصویر ورودی، تصویر احتمال پسین بلوک ایجاد می‌شود که هر پیکسل از این تصویر متناظر با یک بلوک DCT از تصویر ورودی بوده و نمایشگر دست‌کاری بودن آن بلوک از تصویر است.

بعد از ایجاد تصویر احتمال پسین بلوک، این تصویر با اعمال آستانه بهینه  $T_{opt} = \arg \max(\frac{\sigma}{\sigma_0 + \sigma_1})$  به دو ناحیه  $C_0$  و  $C_1$  تقسیم می‌شوند و چهار ویژگی  $T_{opt}$ ،  $\sigma_0$ ،  $\sigma_0 + \sigma_1$  و  $K_0$  استخراج می‌شود که  $K_0$  نمایشگر اتصال در پیکسل‌های دسته  $C_0$  و  $\sigma_0$  و  $\sigma_1$  به ترتیب انحراف معیار تغییرات در دو ناحیه  $C_0$  و  $C_1$  است.

## ۲-۱-۲- ویژگی‌های مبتنی بر توزیع آماری ارقام اول ضرایب DCT

روش بنفورد، روشی آماری است که بر مبنای قانون بنفورد از توزیع‌های احتمالی اولین ارقام اعداد برای صحت‌سنجی آن‌ها استفاده می‌کند. این قانون بیان می‌کند که توزیع احتمال اولین ارقام  $x$  در یک مجموعه از ارقام طبیعی، لگاریتمی هستند. به عبارت دیگر، اگر مجموعه داده‌ای از اعداد از قانون بنفورد تبعیت کند، مهم‌ترین ارقامش (رقم اول از سمت چپ)، توزیع زیر را خواهد داشت:

$$p(x) = \log_{10}\left(1 + \frac{1}{x}\right) \rightarrow x = 1, 2, \dots, 9 \quad (11)$$

که  $p(x)$  احتمال رقم  $x$  است. روش بنفورد برای صحت‌سنجی داده‌ها در کاربردهای متعددی از قبیل حسابداری [۲۵] و اقتصاد [۲۶] استفاده می‌شود. در حوزه پردازش تصویر نیز، توزیع ارقام اول ضرایب بلوک DCT نیز از قانون بنفورد و ضرایب JPEG کوانتیزه‌شده نیز از قانون لگاریتمی مشابه بنفورد پیروی می‌کند. البته این مسئله در صورتی درست است که تصویر دارای یک بار فشرده‌سازی باشد. اما این توزیع به فشرده‌سازی مجدد حساس است، به عبارتی قانون لگاریتمی بعد از دو بار فشرده‌سازی دچار اغتشاش می‌شود. این ویژگی در بسیاری از الگوریتم‌های پردازش تصویر در حوزه جرم‌کاوی و تشخیص جعل نیز استفاده می‌شود [۲۷، ۲۸].

نکته بسیار مهمی که بر عملکرد توزیع آماری ارقام اول ضرایب DCT برای تشخیص جعل تأثیر می‌گذارد انتخاب ضرایب مناسب و نحوه ایجاد توزیع برای ضرایب است. ویژگی‌های مبتنی بر بنفورد، می‌تواند به صورت توزیعی برای کلیه ضرایب DCT در تصویر ایجاد شود که در این حالت بردار ویژگی، شامل ۱۰ مقدار (با در نظر گرفتن مقدار صفر برای ضرایب صفر) برای هر کانال تصویر خواهد بود. در صورتی که توزیع برای هر ضریب DCT به صورت مجزا محاسبه شود

به منظور محاسبه ویژگی‌های مبتنی بر پیروید هیستوگرام ضرایب DCT، باید پیروید برای کلیه ضرایب DCT محاسبه شود. تعداد کل ضرایب DCT برای هر کانال رنگی، ۶۴ ضریب است، که با اندیس‌های (۰ و ۷) تا (۷ و ۷) مشخص می‌شوند. به منظور ایجاد هیستوگرام هر فرکانس یا ضریب DCT با اندیس مشخص، ضریب DCT همه بلوک‌های تصویر با همان اندیس جمع‌آوری و هیستوگرام آن ایجاد می‌شود. یک تصویر رنگی در فشرده‌سازی JPEG، به فضای رنگی YCbCr تبدیل می‌شود. بنابراین می‌توانیم، حداکثر به تعداد ۱۹۲ (۳×۶۴) هیستوگرام، برای ضرایب DCT با فرکانس‌های مختلف و کانال‌های رنگی ایجاد کنیم. ولی از آنجایی که اغلب ضرایب DCT فرکانس بالا، به صفر کوانتیزه می‌شوند، استفاده از آن‌ها مفید نبوده و تنها هیستوگرام‌های فرکانس‌های پایین هر کانال، قابل استفاده می‌باشند. در این تحقیق، با انتخاب سازگار و پویای ضرایب DCT سعی کرده‌ایم که ویژگی‌های مناسبی برای طبقه‌بندی داشته باشیم. ما از سه شیوه متفاوت برای ایجاد هیستوگرام‌ها و محاسبه پیروید استفاده کردیم. که این سه شیوه عبارتند از:

۱- هیستوگرام برای تعدادی از ضرایب فرکانس پایین به ترتیب Zig-Zag تولید می‌شود.

۲- در هر تصویر ورودی، متوسط هر ضریب DCT در هر کانال رنگی محاسبه شده و با اعمال یک حد آستانه مبتنی بر مقدار متوسط، ضرایب با مقادیر بالا و اطلاعات کافی در هر کانال استخراج و هیستوگرام برای آن‌ها محاسبه می‌شود.

۳- در هر تصویر ورودی، متوسط هر ضریب DCT در هر کانال رنگی محاسبه شده و بعد از مرتب‌سازی ضرایب، تعداد  $N_c$  ضریب با مقدار متوسط بالا برای تولید هیستوگرام انتخاب می‌شوند.

به منظور استخراج ویژگی‌های مبتنی بر پیروید هیستوگرام ضرایب DCT، برای هر بلوک  $8 \times 8$  از تصویر، با استفاده از هیستوگرام‌ها و پیرویدهای محاسبه‌شده، احتمال دست‌کاری و یا عدم دست‌کاری بلوک‌ها محاسبه می‌شود. برای این منظور فرض کنید یک پیروید هیستوگرام از اندیس  $s_0$  شروع شده و در اندیس  $(s_0 + p - 1)$  تمام می‌شود. در این حالت احتمال دست‌کاری شدن،  $P_t$ ، و احتمال عدم دست‌کاری،  $P_u$ ، پیکسلی از بلوک که مقدار آن در اندیس  $(s_0 + i)$  از هیستوگرام قرار می‌گیرد به صورت زیر تخمین زده می‌شود [۲۰]:

$$P_u(s_0 + i) = \frac{h(s_0 + i)}{\sum_{k=0}^{p-1} h(s_0 + k)} \quad (7)$$

$$P_t(s_0 + i) = \frac{1}{p} \quad (8)$$

در اینجا،  $h(k)$  مقدار هیستوگرام در بازه یا اندیس  $k$ ام از هیستوگرام ضریب DCT را نشان می‌دهد و  $p$  پیروید هیستوگرام است. همچنین با استفاده از روش نایو بیس، احتمال پسین دست‌کاری یا عدم دست‌کاری آن، به ترتیب به صورت زیر خواهد بود [۲۰]:

تعداد ویژگی‌های مبتنی بر توزیع ارقام اول مستقیماً با تعداد ضرایب و ناحیه فرکانسی انتخاب‌شده ارتباط دارد لذا ما این ویژگی‌ها را برای کلیه ضرایب محاسبه کرده و بعد از تلفیق آن با ویژگی‌های مبتنی بر پیروید هیستوگرام از روش انتخاب ویژگی برای کاهش حجم محاسبات و افزایش صحت طبقه‌بندی استفاده می‌کنیم. این روش علاوه بر انتخاب ناحیه فرکانسی مناسب، امکان انتخاب توزیع رقم با بیشتری تأثیر را نیز فراهم می‌کند. الگوریتم ما برای انتخاب ویژگی یک الگوریتم سلسله مراتبی و شامل دو مرحله است. در مرحله اول الگوریتم انتخاب ویژگی، از معیار F-Score برای انتخاب ویژگی استفاده شده است [۲۹].

معیار F-score معیاری است که میزان تمایز و تفکیک‌پذیری بین دو مجموعه از ویژگی‌ها را اندازه می‌گیرد. اگر  $x_k$  ویژگی  $k$ ام  $(k=1, \dots, m)$ ، از مجموعه بردارهای ویژگی آموزشی باشد و تعداد نمونه‌های دو گروه تصاویر جعلی و غیرجعلی به ترتیب برابر  $n_+$  و  $n_-$  باشند. معیار F-score برای ویژگی  $k$ ام، به صورت زیر تعریف می‌شود [۲۹]:

$$F_{score}(i) = \frac{(x_i^{(+)} - \bar{x}_i)^2 + (x_i^{(-)} - \bar{x}_i)^2}{\frac{1}{n_+ - 1} \sum_{k=1}^{n_+} (x_{k,i}^{(+)} - \bar{x}_i^{(+)})^2 + \frac{1}{n_- - 1} \sum_{k=1}^{n_-} (x_{k,i}^{(-)} - \bar{x}_i^{(-)})^2} \quad (12)$$

که  $x_i^{(-)}$  و  $x_i^{(+)}$  به ترتیب میانگین ویژگی  $k$ ام مجموعه‌های داده کل، جعلی و غیرجعلی می‌باشند.  $x_{k,i}^{(+)}$  و  $x_{k,i}^{(-)}$  نیز به ترتیب ویژگی  $k$ ام برای نمونه  $k$ ام از گروه‌های جعلی و غیرجعلی هستند. مقدار بزرگ معیار F-score برای یک ویژگی نمایانگر تمایز بیشتر آن ویژگی بوده و بر اساس آن، بعد از محاسبه معیار F-score ویژگی‌های با مقدار F-score کم حذف می‌شوند.

از آنجایی که ما برای طبقه‌بندی از ماشین بردار پشتیبان (SVM) استفاده می‌کنیم. لذا در مرحله دوم انتخاب ویژگی، از ضرایب طبقه‌بند ماشین بردار پشتیبان برای انتخاب ویژگی استفاده می‌شود. در روش ماشین بردار پشتیبان برای هر ویژگی یک ضریب محاسبه شده و از تابع زیر برای طبقه‌بندی استفاده می‌شود:

$$f(x) = \text{sgn}(\omega^T \varphi(x) + b) \quad (13)$$

در رابطه فوق  $\text{sgn}$  بیانگر تابع علامت،  $b$  بایاس،  $x$  بیانگر بردار ویژگی،  $\omega$  ضرایب ماشین بردار پشتیبان و تابع  $\phi$  نگاشت‌دهنده بردار ویژگی به فضایی با ابعاد بالاتر هستند. با توجه به رابطه فوق در صورتی که ضریب  $\omega_k$  متناظر با ویژگی  $x_k$  مقدار کمی داشته باشد تأثیر ویژگی  $x_k$  در محاسبه رابطه فوق کم بوده و می‌تواند در مرحله انتخاب ویژگی به‌عنوان یک ویژگی کم‌ارزش حذف شود. بر این اساس ما در مرحله دوم از الگوریتم انتخاب ویژگی، ویژگی‌ها را برحسب معیار  $F_{svm}$  مرتب کرده و ویژگی‌های با مقدار  $F_{svm}$  کم حذف می‌شوند. مقدار  $F_{svm}$  از رابطه زیر تعیین می‌شود [۲۹]:

در این حالت بردار ویژگی، شامل  $640 (10 \times 64)$  مقدار برای هر کانال تصویر خواهد بود. ما در روش پیشنهادی، به‌منظور استخراج بردار ویژگی‌های مبتنی بر توزیع آماری ارقام اول ضرایب DCT از توزیع هر ضریب DCT به‌صورت مجزا استفاده می‌کنیم. الگوریتم پیشنهادی ما برای استخراج این ویژگی‌ها دارای مراحل زیر است:

- ۱- تصویر RGB ناشی از بازگشایی تصویر JPEG به فضای رنگی YCbCr منتقل می‌شود.
- ۲- تصویر هر کانال رنگی به بلوک‌های  $8 \times 8$  غیرهمپوشان تقسیم شده و ضرایب DCT محاسبه می‌شوند.
- ۳- به ازای هر کانال رنگی و هر اندیس از  $64$  ضریب DCT، ضرایب DCT همه بلوک‌های همان کانال و همان اندیس جمع‌آوری و بعد از محاسبه ارقام اول، هیستوگرام آن ایجاد می‌شود.
- ۴- با تجمیع کلیه هیستوگرام‌ها بردار ویژگی هر تصویر ایجاد می‌شود. بردار ویژگی فوق شامل  $1920 (3 \times 10 \times 64)$  مقدار خواهد بود، که یک بردار ویژگی با ابعاد بالا خواهد بود. به‌منظور کاهش حجم محاسبات و نیز انتخاب ویژگی‌های مناسب، ما از یک الگوریتم سلسله مراتبی انتخاب ویژگی استفاده می‌کنیم که در بخش بعد توضیح داده می‌شود.

## ۲-۲- انتخاب ویژگی

همان‌طور که قبلاً ذکر شد انتخاب ناحیه فرکانسی مناسب ضرایب DCT هم بر بازدهی ویژگی‌های مبتنی بر توزیع ارقام اول و هم بر بازدهی ویژگی‌های مبتنی بر پیروید هیستوگرام ضرایب تأثیر به‌سزایی دارد. تعداد ویژگی‌های مبتنی بر هیستوگرام ضرایب  $4$  عدد بوده و تعداد ضرایب انتخابی برای تولید هیستوگرام تأثیری بر تعداد این ویژگی‌ها ندارد. لذا برای انتخاب ناحیه مناسب ضرایب برای ویژگی‌های مبتنی بر پیروید هیستوگرام از روش زیر استفاده شد.

- انتخاب ناحیه‌های فرکانسی متعدد بر اساس شیوه‌های ذکر شده در بخش (۲-۱-۱).
- محاسبه ویژگی‌های مبتنی بر پیروید هیستوگرام
- آموزش ماشین بردار پشتیبان بر اساس ویژگی‌های استخراج‌شده
- طبقه‌بندی تصاویر بر اساس ماشین بردار پشتیبان آموزش داده‌شده
- انتخاب ناحیه‌های فرکانسی با بیشترین درصد طبقه‌بندی صحیح همان‌طور که قبلاً نیز ذکر شد ما از سه شیوه متفاوت برای انتخاب ناحیه‌های فرکانسی و دو شیوه متفاوت برای محاسبه پیروید استفاده کردیم. علاوه‌براین، ما از دو روش مختلف انتخاب کانال رنگی نیز استفاده کردیم که عبارت از استفاده از ۱- فقط اطلاعات کانال Y (اطلاعات سطح خاکستری) و ۲- استفاده از اطلاعات سه کانال رنگی YCbCr هستند. ترکیب این روش‌ها در مجموع ۱۲ روش مختلف را شامل می‌شود و با در نظر گرفتن ۴ ویژگی بهینه برای هر روش در مجموع ۴۸ ویژگی در این مرحله استخراج می‌شود.

دو روش مختلف انتخاب کانال رنگی نیز استفاده کردیم که عبارت از استفاده از ۱- فقط اطلاعات کانال Y (اطلاعات سطح خاکستری) و ۲- استفاده از اطلاعات سه کانال رنگی YCbCr می‌باشد. برای طبقه‌بندی نیز از طبقه‌بند SVM با هسته خطی استفاده شده است. نتایج شکل نشان می‌دهد که استفاده از محاسبه پریود با استفاده از روش مرجع [۲۱] و استفاده از اطلاعات کانال Y نتایج بهتری را به همراه دارد. بهترین نتایج نیز مربوط به استفاده از ۳ ضریب فرکانس پایین است.



(الف)



(ب)

شکل ۳: نمونه‌های از تصاویر آزمایش؛ (الف) تصاویر دست‌کاری نشده، (ب) تصاویر دست‌کاری شده

در شکل ۵، نتایج تشخیص تصاویر جعلی و معتبر با استفاده از روش ایجاد هیستوگرام بر اساس مرتب‌سازی میانگین ضرایب DCT و انتخاب ضرایب بزرگ نشان داده شده است. در این شکل نیز نتایج به ازای دو روش مختلف محاسبه پریود و دو روش مختلف محاسبه انتخاب کانال رنگی نشان داده شده است. نتایج شکل نشان می‌دهد که استفاده از محاسبه پریود با استفاده از روش مرجع [۲۱] و استفاده از اطلاعات کانال Y نتایج بهتری را به همراه دارد.

در شکل ۶ نتایج تشخیص تصاویر جعلی و معتبر با استفاده از روش ایجاد هیستوگرام با اعمال حد آستانه بر روی میانگین ضرایب DCT در هر کانال رنگی نشان داده شده است. در این شکل نیز نتایج به ازای دو روش مختلف محاسبه پریود و دو روش مختلف محاسبه انتخاب کانال رنگی نشان داده شده است. نتایج شکل نشان می‌دهد که استفاده از محاسبه پریود با استفاده از روش مرجع [۲۱] و استفاده از اطلاعات کانال Y نتایج بهتری را به همراه دارد.

بر اساس نتایج شکل‌های ۴ تا ۶ ما از روش‌های زیر برای تولید هیستوگرام و استخراج ۴۸ ویژگی مبتنی بر پریود هیستوگرام استفاده کردیم.

$$F_{svm}(i) \equiv \frac{\bar{\omega}_i^2}{\frac{1}{N_f - 1} \sum_{k=1}^{N_f} (\omega_i - \bar{\omega}_i)^2} \quad (14)$$

در رابطه فوق  $\bar{\omega}_i$  میانگین ضریب ویژگی  $i$  ام در  $N_f$  بار آموزش ماشین بردار پشتیبان بوده و مخرج کسر بیانگر واریانس  $\omega_i$  در  $N_f$  بار آموزش ماشین بردار پشتیبان است. هدف ما در استفاده از رابطه فوق استفاده از ویژگی‌هایی با متوسط بالا و واریانس کم است.

### ۳-۲- طبقه‌بندی

ما برای طبقه‌بندی تصاویر جعلی و غیرجعلی از طبقه‌بند ماشین بردار پشتیبان استفاده کرده‌ایم [۳۰]. ماشین بردار پشتیبان یکی از روش‌های یادگیری با نظارت است که از آن به صورت گسترده برای طبقه‌بندی و رگرسیون استفاده می‌شود. ماشین بردار پشتیبان، داده‌ها را با عبور یک صفحه یا یک ابر صفحه با استفاده از رابطه (۱۳) طبقه‌بندی می‌کند. در این روش ابتدا، نمونه‌هایی که مرز دسته‌ها را تشکیل می‌دهند، محاسبه می‌شوند. به عبارت دیگر، تعدادی از نقاط آموزشی که کمترین فاصله را تا مرز تصمیم‌گیری دارند، می‌توانند به عنوان بردار پشتیبان در نظر گرفته شوند. در این روش، با افزایش ابعاد داده‌ها با هسته‌های مختلف معمولاً نتیجه مطلوب‌تری به دست می‌آید که در قسمت نتایج آزمایشی مورداستفاده قرار گرفته‌اند.

### ۳- نتایج تجربی

به منظور آزمایش و ارزیابی الگوریتم پیشنهادی، این الگوریتم در محیط متلب پیاده‌سازی شد و با پایگاه داده‌ای حاوی ۱۸۳۴ تصویر شامل ۹۱۱ تصویر جعلی و ۹۲۳ تصویر معتبر از پایگاه داده CASIA [۳۱] آزمایش شد. در شکل ۳، نمونه‌هایی از تصاویر پایگاه داده نمایش داده شد. در کلیه آزمایش‌ها ما از ۵۰٪ درصد داده‌ها برای آموزش و از ۵۰٪ باقی‌مانده برای آزمایش استفاده کرده‌ایم. همچنین مراحل آموزش و آزمایش ۱۰ مرتبه تکرار شده و متوسط نتایج به عنوان نتیجه نهایی ذکر شده است.

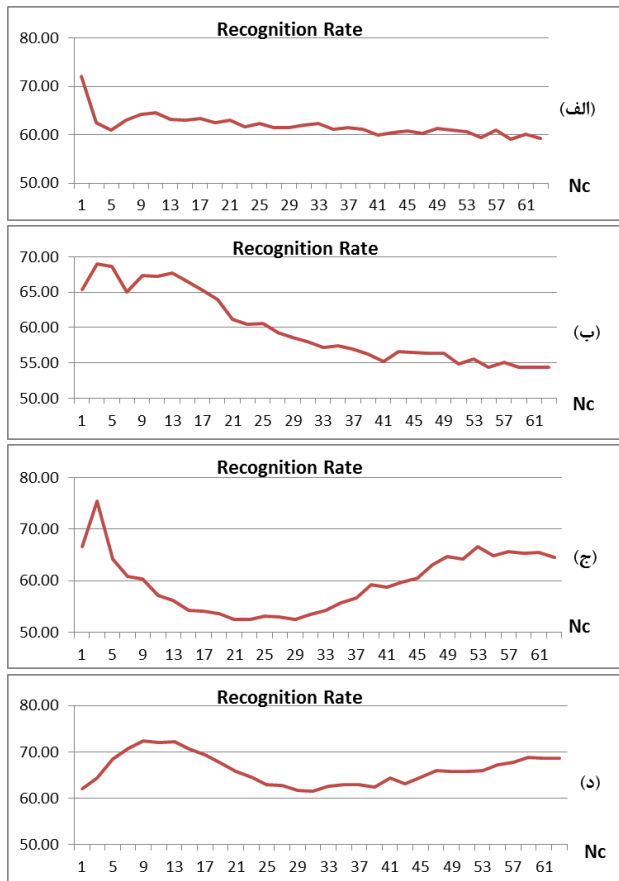
همان‌طور که ذکر شد، ما از سه روش مختلف برای انتخاب ضرایب DCT مناسب برای استخراج ویژگی‌های مبتنی بر پریود هیستوگرام استفاده کردیم که عبارتند از:

۱- ایجاد هیستوگرام برای تعدادی از ضرایب فرکانس پایین به ترتیب Zig-Zag

۲- مرتب‌سازی میانگین ضرایب DCT و انتخاب ضرایب بزرگ.

۳- اعمال حد آستانه بر روی میانگین ضرایب DCT در هر کانال رنگی

در شکل ۴ نتایج تشخیص تصاویر جعلی و معتبر با استفاده از روش ایجاد هیستوگرام برای تعدادی از ضرایب فرکانس پایین به ترتیب Zig-Zag به ازای تعداد ضرایب مختلف نشان داده شده است. در این شکل، نتایج به ازای دو روش مختلف محاسبه پریود نشان داده شده است. علاوه بر انتخاب دو روش مختلف برای محاسبه پریود، ما از

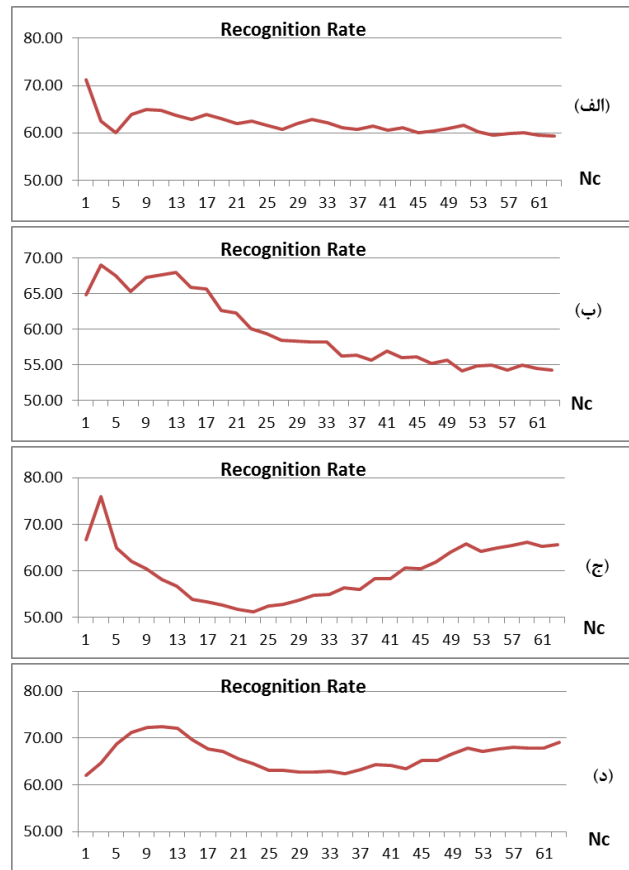
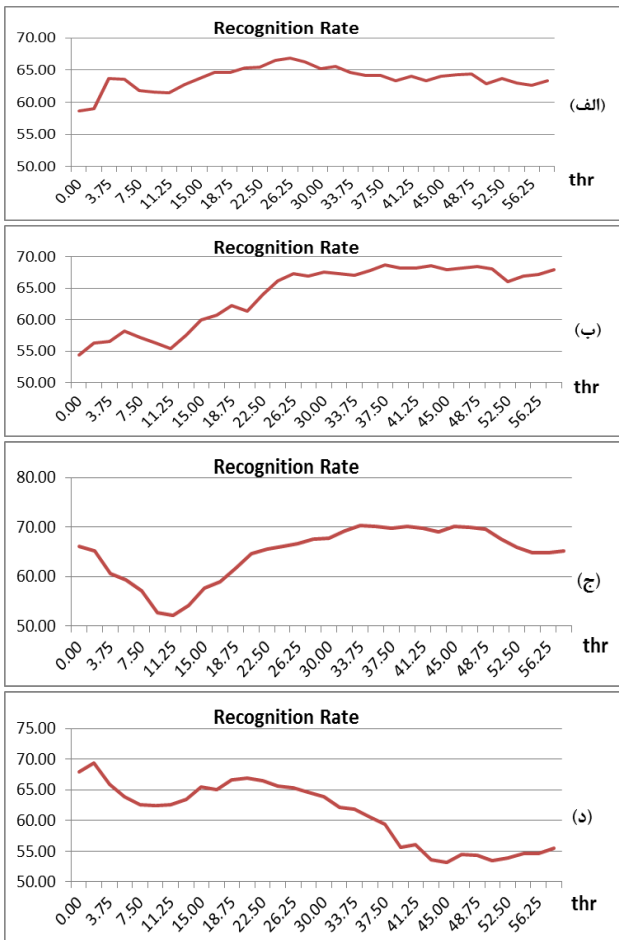


شکل ۴: نتایج تشخیص تصاویر جعلی و معتبر با استفاده از روش ایجاد هیستوگرام برای تعدادی از ضرایب فرکانس پایین (Nc) به ترتیب Zig-Zag؛ (الف) روش پریود مرجع [۲۰] و کانال Y، (ب) روش پریود مرجع [۲۰] و کانال‌های YCbCr، (ج) روش پریود مرجع [۲۱] و کانال Y، (د) روش پریود مرجع [۲۱] و کانال‌های YCbCr

در شکل ۷ و جدول ۱ ما از ۵۰٪ درصد داده‌ها برای آموزش و از ۵۰٪ باقی‌مانده برای آزمایش استفاده کرده‌ایم. همچنین مراحل آموزش و آزمایش ۱۰ مرتبه تکرار شده و متوسط نتایج به‌عنوان نتیجه نهایی ذکر شده است. بر اساس نتایج جدول ۱ بهترین نتایج مربوط به SVM خطی با  $N_{SVM}=450$  و  $N_{FScore}=1400$  است. در این حالت الگوریتم پیشنهادی دارای صحت شناسایی ۹۹/۵۵ درصد است. به‌منظور مقایسه عملکرد الگوریتم پیشنهادی با سایر روش‌های موجود ما الگوریتم‌های مراجع [۲۰]، [۲۱] و [۲۲] را پیاده‌سازی و با پایگاه داده توضیح داده‌شده مورد آزمایش قرار دادیم. در این آزمایش‌ها نیز از ۵۰٪ درصد داده‌ها برای آموزش و از ۵۰٪ باقی‌مانده برای آزمایش استفاده کرده‌ایم. همچنین مراحل آموزش و آزمایش ۱۰ مرتبه تکرار شده و متوسط نتایج به‌عنوان نتیجه نهایی ذکر شده است. جدول ۲ نتایج الگوریتم پیشنهادی را با روش‌های موجود پیاده‌سازی شده مقایسه می‌کند. همان‌طور که جدول ۲ نشان می‌دهد، الگوریتم پیشنهادی توانسته به میزان ۷/۶ درصد نتیجه شناسایی را بهبود بخشد.

- استفاده از یک ضریب فرکانس پایین به ترتیب zigzag در کانال Y و روش محاسبه پریود مرجع [۲۰].
  - استفاده از ۳ ضریب فرکانس پایین به ترتیب zigzag در کانال‌های YCbCr و روش محاسبه پریود مرجع [۲۰].
  - استفاده از ۳ ضریب فرکانس پایین به ترتیب zigzag در کانال Y و روش محاسبه پریود مرجع [۲۱].
  - استفاده از ۹ ضریب فرکانس پایین به ترتیب zigzag در کانال‌های YCbCr و روش محاسبه پریود مرجع [۲۱].
  - استفاده از یک ضریب DCT با دامنه بزرگ و بهره‌گیری از کانال Y و روش محاسبه پریود مرجع [۲۰].
  - استفاده از ۳ ضریب DCT با دامنه بزرگ و بهره‌گیری از کانال‌های YCbCr و روش محاسبه پریود مرجع [۲۰].
  - استفاده از ۳ ضریب DCT با دامنه بزرگ و بهره‌گیری از کانال Y و روش محاسبه پریود مرجع [۲۱].
  - استفاده از ۱۱ ضریب DCT با دامنه بزرگ و بهره‌گیری از کانال‌های YCbCr و روش محاسبه پریود مرجع [۲۱].
  - استفاده از حد آستانه ۲۶/۲۵ به متوسط ضرایب DCT و بهره‌گیری از اطلاعات کانال Y و روش محاسبه پریود مرجع [۲۰].
  - استفاده از حد آستانه ۳۷/۵ به متوسط ضرایب DCT و بهره‌گیری از اطلاعات کانال‌های YCbCr و روش محاسبه پریود مرجع [۲۰].
  - استفاده از حد آستانه ۳۳/۷۵ به متوسط ضرایب DCT و بهره‌گیری از اطلاعات کانال Y و روش محاسبه پریود مرجع [۲۱].
  - استفاده از حد آستانه ۱/۸۸ به متوسط ضرایب DCT و بهره‌گیری از اطلاعات کانال‌های YCbCr و روش محاسبه پریود مرجع [۲۱].
- بعد از انتخاب ویژگی‌های مبتنی بر پریود هیستوگرام این ویژگی‌ها با ویژگی‌های مبتنی بر توزیع آماری ارقام اول ضرایب DCT تلفیق شده و یک ویژگی ۱۹۶۸ بعدی ایجاد می‌شود. همان‌طور که ذکر شد به‌منظور کاهش حجم محاسبات و افزایش صحت ما از روش انتخاب ویژگی استفاده می‌کنیم. روش انتخاب ویژگی پیشنهادی، یک الگوریتم دومرحله‌ای است. در مرحله اول با استفاده از معیار F-score، به تعداد  $N_{FScore}$  از ویژگی‌ها انتخاب و در مرحله دوم با استفاده از روش SVM تعداد ویژگی‌ها به  $N_{SVM}$  کاهش می‌یابد. شکل ۷ درصد شناسایی تصاویر جعلی و معتبر را به ازای مقادیر مختلف  $N_{FScore}$  و  $N_{SVM}$  را نشان می‌دهد. در این شکل درصد شناسایی با استفاده از SVM با هسته خطی و RBF نشان داده شده است. در جدول ۱ حداکثر درصد شناسایی متوسط تصاویر جعلی و معتبر به ازای مقادیر مختلف  $N_{FScore}$  و  $N_{SVM}$  با SVM با هسته‌های خطی، RBF، Quadratic و MLP نشان داده شده است.





شکل ۶: نتایج تشخیص تصاویر جعلی و معتبر با استفاده از روش ایجاد هیستوگرام با اعمال حد آستانه (thr) بر روی میانگین ضرایب DCT: (الف) روش پریود مرجع [۲۰] و کانال Y، (ب) روش پریود مرجع [۲۰] و کانال‌های YCbCr، (ج) روش پریود مرجع [۲۱] و کانال Y، (د) روش پریود مرجع [۲۱] و کانال‌های YCbCr

شکل ۵: نتایج تشخیص تصاویر جعلی و معتبر با استفاده از روش ایجاد هیستوگرام برای تعدادی از ضرایب (Nc) بر اساس مرتب‌سازی میانگین ضرایب DCT و انتخاب ضرایب بزرگ؛ (الف) روش پریود مرجع [۲۰] و کانال Y، (ب) روش پریود مرجع [۲۰] و کانال‌های YCbCr، (ج) روش پریود مرجع [۲۱] و کانال Y، (د) روش پریود مرجع [۲۱] و کانال‌های YCbCr

جدول ۱: حداکثر درصد شناسایی متوسط تصاویر جعلی و معتبر به

ازای هسته‌های مختلف SVM

هسته	مقدار $N_{SVM}$	مقدار $N_{Fscore}$	در صد شناسایی
خطی	۴۵۰	۱۴۰۰	۹۹/۵۵
RBF	۲۰	۵۰۰	۹۷/۶۹
Quadratic	۳۲۰	۶۰۰	۹۹/۲۰
MLP	۱۷۰	۱۵۰۰	۸۷/۶۴

جدول ۲: مقایسه درصد شناسایی الگوریتم پیشنهادی با سایر

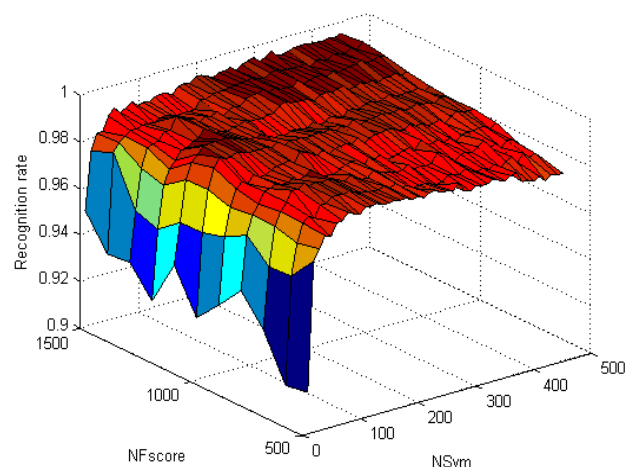
روش‌های موجود

روش	الگوریتم پیشنهادی	روش مرجع [۲۰]	روش مرجع [۲۱]	روش مرجع [۲۷]
درصد شناسایی	۹۹/۵۵	۷۲/۰۷	۷۵/۴۴	۹۱/۹۵

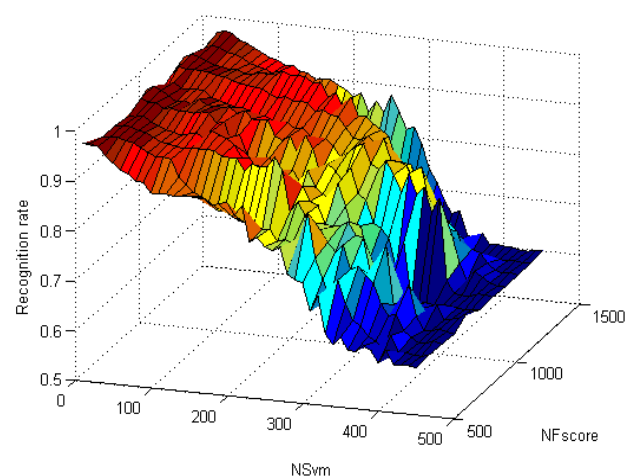
#### ۴- نتیجه‌گیری

در این مقاله، روش جدیدی برای شناسایی تصاویر دیجیتال جعلی با استفاده از تشخیص فشرده‌سازی مجدد و ویژگی‌های مبتنی بر اثرات کوانتیزاسیون ضرایب DCT شامل ویژگی‌های پریود کوانتیزاسیون و توزیع ارقام اول ضرایب DCT ارائه شد. در الگوریتم پیشنهادی از یک الگوریتم سلسله مراتبی و جدید برای انتخاب ویژگی استفاده شد و الگوریتم پیشنهادی با پایگاه داده‌ای حاوی ۱۸۳۴ تصویر شامل ۹۱۱ تصویر جعلی و ۹۲۳ تصویر معتبر از پایگاه داده CASIA آزمایش شد. نتایج آزمایش‌ها نشانگر صحت شناسایی ۹۹/۵۵ درصد و بهبود ۷/۶ درصد نسبت به روش‌های موجود است.

- IEEE International Conference on Image Processing (ICIP)*, 2007.
- [5] M. K. Johnson, and H. Farid, "Exposing digital forgeries through chromatic aberration," *8th Workshop on Multimedia and Security*, pp. 48-55, 2006.
- [6] H. Cao, and A. C. Kot, "Accurate detection of demosaicing regularity for digital image forensics," *IEEE Transactions on Information Forensics and Security*, vol. 4, no. 4, pp. 899-910, 2009.
- [7] A. C. Popescu, and H. Farid, "Exposing digital forgeries in color filter array interpolated images," *IEEE Transactions on Signal Processing*, vol. 53, no. 10, pp. 3948-3959, 2005.
- [8] I. Amerini, L. Ballan, R. Caldelli, A. Del Bimbo, L. Del Tongo, and G. Serra, "Copy-move forgery detection and localization by means of robust clustering with J-Linkage," *Signal Processing: Image Communication*, vol. 28, no. 6, pp. 659-669, 2013.
- [9] I. Amerini, L. Ballan, R. Caldelli, A. Del Bimbo, and G. Serra, "A sift-based forensic method for copy-move attack detection and transformation recovery," *IEEE Transactions on Information Forensics and Security*, vol. 6, no. 3, pp. 1099-1110, 2011.
- [10] S. D. Lin, and T. Wu, "An integrated technique for splicing and copy-move forgery image detection," *4th International Congress on Image and Signal Processing (CISP)*, pp. 1086-1090, 2011.
- [11] M. Minati, and M. Adhikary, "Detection of clones in digital images," *International Journal of Computer Science*, vol. 9, no. 1, pp. 91-102, 2014.
- [12] A. C. Popescu, and H. Farid, *Exposing digital forgeries by detecting duplicated image regions*, Dept. Comput. Sci., Dartmouth College, Tech. Rep. TR2004-515, 2004.
- [13] A. Kashyap, B. Suresh, M. Agrawal, H. Gupta, and S. D. Joshi, "Detection of splicing forgery using wavelet decomposition," *International Conference on Computing, Communication & Automation (ICCCA)*, pp. 843-848, 2015.
- [14] E. Kee, J. O'Brien, and H. Farid, "Exposing photo manipulation with inconsistent shadows," *ACM Transactions on Graphics (TOG)*, vol. 32, no. 3, pp. 28, 2013.
- [15] H. Farid, and M. J. Bravo, "Image forensic analyses that elude the human visual system," *Proc. SPIE 7541, Media Forensics and Security II*, 2010.
- [16] L. Wu, and Y. Wang, "Detecting image forgeries using geometric cues," *Computer Vision for Multimedia Applications: Methods and Solutions*, pp. 197-217, 2010.
- [17] V. Conotter, J. F. O'Brien, and H. Farid, "Exposing digital forgeries in ballistic motion," *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 1, pp. 283-296, 2012.
- [18] L. Bin, N. Tian-Tsong, L. XiaoLong, T. Shunquan, and H. Jiwu, "Revealing the trace of high-quality JPEG compression through quantization noise analysis," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 3, pp. 558-573, 2015.



(الف)



(ب)

شکل ۷: درصد شناسایی تصاویر جعلی و معتبر را به ازای مقادیر مختلف NScore و NSVM؛ (الف) استفاده از هسته خطی، (ب) استفاده از هسته RBF.

## مراجع

- [۱] مرتضی نصیری و علیرضا بهراد، «استخراج و بازسازی خودکار اطلاعات سه‌بعدی زمین فوتبال در تصاویر تک‌دوربینی کالیبره نشده»، *مجله مهندسی برق دانشگاه تبریز*، جلد ۴۴، شماره ۴ صفحات ۵۵-۶۴، ۱۳۹۳.
- [۲] طناز اکبریور و سیلان دانشور، «افزایش کیفیت ادغام تصاویر MRI و PET با استفاده از ترکیب موجک گابور و مدل شبکه‌ای»، *مجله مهندسی برق دانشگاه تبریز*، جلد ۴۴، شماره ۴، صفحات ۲۳-۳۵، ۱۳۹۴.
- [3] J. Lukas, J. Fridrich, and M. Goljan, "Digital camera identification from sensor pattern noise," *IEEE Transactions on Information Forensics and Security*, vol. 1, no. 2, pp. 205-214, 2006.
- [4] A. E. Dirik, S. Bayram, H. T. Sencar, and N. Memon, "New features to identify computer generated images,"

- [26] K. H. Tödter, "Benford's law as an indicator of fraud in economics," *German Economic Review*, vol. 10, no. 3, pp. 339-351, 2009.
- [27] S. Milani, M. Tagliasacchi, and S. Tubaro, "Discriminating multiple JPEG compressions using first digit features," *APSIPA Transactions on Signal and Information Processing*, vol. 3, pp. e19, 2014.
- [28] X. Zhao, A. T. Ho, and Y. Q. Shi, "Image forensics using generalised Benford's law for accurate detection of unknown JPEG compression in watermarked images," *16th International Conference on Digital Signal Processing*, pp. 1-8, 2009.
- [29] Y. W. Chen, and C. J. Lin, "Combining SVMs with various feature selection strategies," *Feature extraction*, I. Guyon, M. Nikravesh, S. Gunn and L. A. Zadeh, eds., pp. 315-324, Springer, 2006.
- [30] C. Cortes, and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273-297, 1995.
- [31] D. Jing, W. Wei, and T. Tieniu, "CASIA image tampering detection evaluation database," *IEEE China Summit & International Conference on Signal and Information Processing (ChinaSIP)*, pp. 422-426, 2013.
- [19] A. Popescu, and H. Farid, "Statistical tools for digital forensics," *Information Hiding*, J. Fridrich, ed., pp. 128-147, Springer, 2005.
- [20] Z. Lin, J. He, X. Tang, and C. K. Tang, "Fast, automatic and fine-grained tampered JPEG image detection via DCT coefficient analysis," *Pattern Recognition*, vol. 42, no. 11, pp. 2492-2501, 2009.
- [21] V. L. Thing, Y. Chen, and C. Cheh, "An improved double compression detection method for JPEG image forensics," *IEEE International Symposium on Multimedia*, pp. 290-297, 2012.
- [22] B. Li, Y. Q. Shi, and J. Huang, "Detecting doubly compressed JPEG images by using mode based first digit features," *IEEE 10th Workshop on Multimedia Signal Processing*, pp. 730-735, 2008.
- [23] S. Milani, M. Tagliasacchi, and S. Tubaro, "Discriminating multiple jpeg compression using first digit features," *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 2253-2256, 2012.
- [24] C. Pasquini, F. Perez-Gonzalez, and G. Boato, "A Benford-Fourier JPEG compression detector," *IEEE International Conference on Image Processing (ICIP)*, pp. 5322-5326, 2014.
- [25] C. Durtschi, W. Hillison, and C. Pacini, "The effective use of Benford's law to assist in detecting fraud in accounting data," *Journal of Forensic Accounting*, vol. 5, no. 1, pp. 17-34, 2004.

## زیر نویس ها

- 
- <sup>1</sup> Watermarking  
<sup>2</sup> Photo Response Non Uniformity  
<sup>3</sup> Splicing  
<sup>4</sup> Joint Photographic Experts Group  
<sup>5</sup> Discrete Cosine Transform  
<sup>6</sup> Offline  
<sup>7</sup> Naïve Bayesian  
<sup>8</sup> Support Vector Machine