

مقاله تقسیم بندی تصویر بر مبنای مدل های فضایی رنگی مختلف با استفاده از

GrabCut اتوماتیک

این مقاله یک مطالعه مقایسه ای با استفاده از فضاهای مختلف رنگی برای ارزیابی عملکرد تقسیم بندی تصاویر رنگی با استفاده از تکنیک GrabCut اتوماتیک ارائه می دهد. اتوماسیون روش GrabCut به عنوان یک اصلاح نیمه اتوماتیک اصلی برای از بین بردن تعامل کاربر پیشنهاد شده است. GrabCut خودکار به صورت غیر نظارت شده از تکنیک خوشه بندی درختی و بومان برای مرحله قالب بندی (فرمت) GrabCut خودکار که با فضای رنگی RGB ، HSV ، CMY ، XYZ و YUV اعمال می شود، استفاده می کند. مطالعه تطبیقی و نتایج تجربی با استفاده از تصاویر رنگی مختلف نشان می دهد که فضای رنگی RGB بهترین نمایش فضای رنگی است که برای مجموعه ای از تصاویر استفاده شده است.

1- معرفی

فرآیند تقسیم یک تصویر دیجیتال به چندین بخش ها به عنوان تقسیم بندی تصویر تعریف می شود. تقسیم بندی هدف آن تقسیم تصویر به مناطقی است که می توانند بیشتر نمایانگر باشد و تجزیه و تحلیل آن آسان تر است. چنین مناطقی ممکن است با سطوح فردی، اشیاء، یا قسمت های طبیعی اشیاء مطابقت داشته باشد. به طور معمول تقسیم بندی تصویر فرایندی است که برای تعیین محل اشیاء و مرزها در تصاویر استفاده می شود (به عنوان مثال، خطوط یا منحنی ها). علاوه بر این، می توان آن را به عنوان فرآیند برچسب زدن هر پیکسل در یک تصویر تعریف کرد، که در آن تمام پیکسل هایی که دارای برچسب مشابه هستند برخی از ویژگی های بصری را به اشتراک می گذارند. معمولاً تقسیم بندی برای محاسبه بهترین تقسیم بندی از اطلاعات محلی در تصویر دیجیتال استفاده می

کند، مانند اطلاعات رنگ که برای ایجاد هیستوگرام یا اطلاعات نشان دهنده لبه ها، مرزها، یا اطلاعات بافت استفاده می شود.

تقسیم تصویر رنگی که بر اساس ویژگی رنگ پیکسل های تصویر است فرض می کند که رنگ های همگن در تصویر مربوط به خوشه های جداگانه است و از این رو به معنی اشیاء در آن تصویر است به عبارت دیگر، هر خوشه ای یک کلاس از پیکسل ها را تعریف می کند که خصوصیات مشابه رنگ را به اشتراک می گذارند. همانطور که نتایج تقسیم بندی به فضای رنگ مورد استفاده بستگی دارد فضای رنگ تنهایی وجود ندارد که بتواند نتایج قابل قبول را برای انواع تصاویر فراهم کند. به همین دلیل، بسیاری از نویسندگان سعی کردند فضای رنگی را مشخص کنند که مشکل تقسیم تصویر رنگی خاص آنها را انطباق خواهد داد. در این کار، تقسیم بندی تصاویر رنگی با فضاهای رنگی کلاسیک مختلف، *RGB*، *CMY*، *XYZ*، *YUV*، و *HSV* آزمایش شده است که انتخاب بهترین فضای رنگی برای انواع تصاویر در نظر گرفته شده است.

روند تقسیم بندی بر اساس GrabCut تکنیک تقسیم بندی است که با اندیشه صحیح به عنوان یکی از تکنیک های وضعیت هنری قدرتمند در مورد مشکل تقسیم بندی تصویر رنگی مطرح شده است. طرح به حداقل رساندن انرژی تکراری GrabCut بر اساس بهینه سازی قدرتمند روش برش گراف است که اجازه می دهد تقسیم بهینه جهانی گسترش پیدا کند. علاوه بر این، نمودار برش می تواند به راحتی مشکل تصاویر N-D را توسعه دهد. علاوه بر این، کاهش هزینه عملکرد انرژی فرایند برش نمودار اجازه می دهد تا آن را از نظر ویژگی های تصویری مختلف مانند رنگ، منطقه، مرز، یا هر گونه مخلوط از ویژگی های تصویری تعریف کند. این انعطاف پذیری پتانسیل وسیعی برای استفاده از GrabCut در برنامه های مختلف فراهم می کند. از سوی دیگر، GrabCut به عنوان یک تکنیک تقسیم بندی دولبه محسوب می شود، که در آن تصاویر به دو منطقه پس زمینه و پیش زمینه فقط می توانند تقسیم شوند. مداخله کاربر اولیه برای مشخص کردن یک شی مورد نظر برای جدا شدن از تصویر، با توجه به همه پیکسل های تصویر باقی مانده به عنوان یک منطقه پس زمینه به درستی مورد نیاز است. این GrabCut را به عنوان روش تقسیم بندی نیمه اتوماتیک طبقه بندی می کند و کیفیت اولیه را تغییر می دهد و در نتیجه عملکرد تقسیم بندی، حساس به انتخاب کاربر می شود. به عبارت دیگر، مقداردهی اولیه GrabCut ضعیف ممکن است منجر به

دقت تقسیم بندی نهایی بد شود که ممکن است به تعاملات کاربر اضافی با این نتایج تقسیم بندی برای تنظیم خوب نیاز داشته باشد.

در این کار GrabCut اصلاح شده به صورت تکنیک تقسیم بندی اتوماتیک ارائه می شود ، که می تواند تصویر را به اشیاء طبیعی بدون نیاز به مداخله کاربر اولیه جدا کند. اتوماسیون GrabCut با استفاده از خوشه بندی درختی و بومان به عنوان خوشه بندی غیر نظارت شده اعمال می شود. انتخاب خوشه بندی درختی و بومان بر اساس نتایج مقایسه تجربی در این کار انجام شد. این مقاله از استفاده برخی از معیارهای ارزیابی قدرت تفکیک کننده از GrabCut اتوماتیک با فضاهای رنگی مختلف بهره برداری شده است. ادامه مقاله به شرح زیر تدوین شده است. بخش 2 یک پس زمینه پایه ای روی تقسیم بندی بر پایه مدل های فضایی رنگی ، تقسیم بندی تصویر با استفاده از GrabCut و تکنیک های خوشه بندی بدون نظارت فراهم شده است. بخش 3 مدل های فضای رنگی مختلف را توضیح می دهد بخش 4 با مثال خوشه بندی درختی و بومان را نشان می دهد . تکنیک اصلی GrabCut و جزئیات اصلاح آن در بخش 5 توضیح داده شده است. نتایج تجربی در بخش 6 ارائه شده است، در حالی که نتیجه گیری و کار آتی در بخش 7 ارائه شده است.

2- کار مرتبط

همانطور که هیچ عقیده رایجی در مورد اینکه بهترین انتخاب برای فضای رنگی بر پایه تقسیم تصویر است بیان نشده است، برخی کارهای تحقیقاتی برای شناسایی بهترین فضای رنگی برای یک وظیفه خاص تلاش کردند چندین کار نشان می دهد که فضاهای رنگی مختلف برای مسئله تقسیم تصویر رنگی مفید است. یوریو و همکاران یک مطالعه تطبیقی بین فضاهای رنگی مختلف در تقسیم بندی تصویر بر اساس خوشه با استفاده از دو الگوریتم خوشه بندی مشابه انجام داده اند. مطالعه مورد بحث آنها تستی از چهار فضای رنگی، RGB ، CMY ، HSV ، و YUV ، به منظور شناسایی بهترین نمایش رنگ بود. آنها در بیشتر موارد با استفاده از فضای رنگی CMY بهترین نتایج خود را به دست آوردند، در حالی که HSV همچنین نتایج خوبی ارائه داد. بوشین و همکاران یک روش برای انتخاب اتوماتیک یک فضای رنگی خاص در میان فضاهای رنگی کلاسیک پیشنهاد دادند. این انتخاب با توجه به معیار ارزیابی براساس یک آنالیز رنگ طیفی بود. این معیار کیفیت تقسیم بندی در هر فضا و انتخاب بهترین را

ارزیابی می کند ، که خواص خاص خود را حفظ می کند. یک مطالعه روی بیش از ده نفر در فضای رنگ مشترک برای تشخیص رنگ پوست ارائه شده است. آنها نتیجه گرفتند که HSV بهترین فضای رنگ برای تشخیص پوست در یک تصویر است. مطالعه دیگری که برای طبقه بندی توپینگ پیتزا مورد استفاده قرار گرفت ثابت کرد که طبقه بندی چندجمله ای SVM همراه با فضای رنگی HSV بهترین روش در میان پنج فضای رنگ مختلف است. برپایه یک مطالعه تطبیقی بین مدل RGB و HSV روییز و همکاران اعلام کردند که بهترین دقت با نمایش HSV بدرستی برای رسیدن به پردازش زمانی واقعی در زمین های زراعی واقعی برای تقسیم بندی محصول بدست آمده بود.

$GrabCut$ یکی از تکنیک های قدرتمند محسوب می شود که برای تقسیم تصویر رنگی استفاده می شود. این برای مشکلات تقسیم بندی مختلف مانند تقسیم بندی بدن انسان، تقسیم بندی ویدئو، تقسیم بندی معنایی و تقسیم بندی حجم مورد استفاده قرار میگیرد. در استخراج خودکار بدن انسان از تصاویر رنگی رشد یافته بود توسط hu . تکرار تکنیک $GrabCut$ برای آپدیت پویا یک شکل $trimap$ استفاده شده بود، که از نتایج یک آشکارساز اسکن شده برای تشخیص چهره از تصاویر مقداردهی اولیه شده استفاده شده بود. این تحقیقات یسری اشکالاتی دارد مانند فرآیندهای چند مرحله ای و تکراری، علاوه بر اینکه محدود به ژستهای انسان است با چهره های جانبی جلو است. متولوژی تقسیم بندی انسانی $GrabCut$ فضایی کاملاً اتوماتیک توسط هرناندز و همکاران پیشنهاد شد. آنها روشی را توسعه دادند که مزایای ترکیبی از پیگردی و تقسیم بندی را داشت. به جای مداخله کاربر اولیه در مقداردهی اولیه الگوریتم $GrabCut$ ، مجموعه ای از دانه های تعریف شده توسط تشخیص صورت و یک مدل رنگ پوست برای مقداردهی اولیه استفاده می شود. یک رویکرد دیگر برای تقسیم انسانها از تصاویر درهم و برهم توسط گلشن و همکاران پیشنهاد شد. آنها مدل رنگی محلی بر مبنای $GrabCut$ برای تقسیم بندی اتوماتیک به کار بردند. این مدل رنگی محلی $GrabCut$ برای اصلاح تقسیم بندی های انسانی خام که آنها بدست آوردند مورد استفاده قرار گرفت. در تقسیم ویدئو $Corrigan$ و همکاران $GrabCut$ را برای تقسیم بندی شیئی ویدیوهای قوی تر توسعه دادند. آنها مدل ترکیبی $Gaussian$ (GMM) از الگوریتم $GrabCut$ را گسترش دادند، به طوری که فضای رنگی با مشتق زمان از شدت پیکسل به منظور شامل شدن اطلاعات زمانی در روند بهینه سازی تقسیم بندی کامل شده بود. $G\ddot{o}ring$ و همکاران $GrabCut$ را به یک چارچوب تقسیم بندی معنایی با برجسب گذاری اشیاء در یک تصویر داده شده متصل کردند. اخیراً، رامیرزو همکاران یک طرح کاملاً موازی با استفاده از $GrabCut$

پیشنهاد کرد برای تقسیم بندی سه بعدی که برای اجرا روی GPU پذیرفته شده است. این طرح با هدف تولید نتایج تقسیم بندی برای مورد شبکه های حجمی، علاوه بر این کاهش زمان محاسبات کارآمد می باشد.

خوشه بندی، طبقه بندی غیر نظارت شده الگوها به گروه ها، یکی از مهم ترین وظایف در تجزیه و تحلیل اطلاعات اکتشافی است. این یک تاریخ طولانی و غنی در بسیاری از رشته های علمی از جمله انسان شناسی، زیست شناسی، پزشکی، روانشناسی، آمار، ریاضیات، مهندسی، و علوم کامپیوتر است. خوشه بندی در تقسیم بندی های تصویر به عنوان فرآیند شناسایی گروه های اولیه تصاویر مشابه تعریف می شود. تکنیک های خوشه بندی غیر نظارت شده خوشه بندی مبتنی بر محتوا هستند، که در آن محتوا به اشکال، بافت ها و یا هر گونه اطلاعات دیگر که از تصویر خودش به ارث برده می تواند اشاره می کند .

در این موارد از تقسیم بندی دو تکه، تفکیک خوب بین پیش زمینه و پس زمینه مورد نیاز است. و می تواند از طریق یافتن خوشه ها با واریانس کم انجام شود، از آنجا که خوشه آسان تر از بقیه جدا می شود. انتخاب روش خوشه بندی درختی و بومان توسط روسون و توماسی و چاونگ و همکاران برای به دست آوردن خوشه های جدا شده محکم و خوب هدایت می شود. آنها برای حل مشکل مات کردن (شطرنجی کردن) تصویر کار کرده اند که برای ترکیب تصویر مورد نیاز است. در رویکرد خود، الگوریتم تقسیم باینری درختی و بومان برای پارتیشن بندی رنگ های منطقه ناشناخته به چند دسته استفاده شده است و برای تولید یک توزیع رنگی برای منطقه ناشناخته تخمین زده می شود. بر اساس یک مطالعه مقایسه ای در خوشه بندی درختی و بومان از دیگر تکنیک های خوشه بندی بدون نظارت از جمله خود سازماندهی نقشه ها (SOFM) و فازی C-means (FCM) برای اتوماسیون از GrabCut با توجه به بهبود دقت تقسیم بندی برتر است.

3-مدل های فضای رنگی

فضای رنگی که به طور گسترده ای استفاده می شود فضای رنگی RGB است، از آنجا که فضاهای رنگی زیادی وجود دارد ، بهتر است که آنها را به دسته های کمتر با توجه به تعاریف و خواص آنها طبقه بندی کنیم. Vandenbroucke طبقه بندی رنگ را به مقوله های زیر پیشنهاد کرد:

(i) فضاهای اولیه که بر اساس این نظریه استوار است: فضاهای اولیه RGB واقعی، CMY ضعیف، و فضاهای

اولیه XYZ خیالی

تبدیل از RGB به CMY :

$$C' = 1 - R \quad C = \min(1, \max(0, C' - K'))$$

$$M' = 1 - G \quad M = \min(1, \max(0, M' - K')) \quad (1)$$

$$Y' = 1 - B \quad Y = \min(1, \max(0, Y' - K'))$$

$$K' = \min(C', M', Y')$$

و تبدیل از RGB به XYZ :

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.412453 & 0.357580 & 0.180423 \\ 0.212671 & 0.715160 & 0.072169 \\ 0.019334 & 0.119193 & 0.950227 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (2)$$

(ا) فضای رنگ YUV که مثالی از فضاهای رنگ تابی روشنایی است. تبدیل از RGB به YUV :

$$\begin{bmatrix} Y \\ U \\ V \end{bmatrix} = \begin{bmatrix} 0.2989 & 0.5866 & 0.1145 \\ -0.147 & -0.289 & 0.436 \\ 0.615 & -0.515 & -0.100 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (3)$$

(II) فضاهای ادراکی که سعی در اندازه گیری ذهنی درک رنگ انسان از طریق معنای سه اندازه، شدت، رنگ و اشباع دارد: HSV یک نمونه از فضای رنگ ادراکی است. این تبدیل از RGB به HSV است:

$$\begin{aligned}
 & \left[\begin{aligned} & 0, & \text{if Max=Min} \\ & 60^\circ \times (G-B/\text{Max}-\text{Min}) + 360^\circ, & \end{aligned} \right. \\
 & \left[\begin{aligned} & H = \quad \times \text{Mod } 360^\circ, & \text{if Max=R} \\ & 60^\circ \times (B-R/\text{Max}-\text{Min}) + 120^\circ, & \text{if Max=G} \\ & 60^\circ \times (R-G/\text{Max}-\text{Min}) + 240^\circ, & \text{if Max=B} \end{aligned} \right. \quad (4) \\
 & \left[\begin{aligned} & 0, & \text{if Max=0} \\ & S = \quad \text{Max}-\text{Min}/\text{Max} & \text{Otherwise} \end{aligned} \right. \\
 & V = \text{Max}.
 \end{aligned}$$

4- تکنیک خوشه بندی درختی و بومان

درختی و بومان یک تکنیک خوشه بندی کوانتومی رنگ است که برای تعیین خوشه خوب تقسیم شده از بردار مشخصه ماتریس کوواریانس رنگ استفاده می کند. این الگوریتم با تمام پیکسل ها در یک خوشه تکی شروع می شود. این خوشه سپس با استفاده از یک تابع بردار مشخصه از ماتریس کوواریانس به عنوان نقطه تقسیم به دو قسمت تقسیم می شود. سپس از مقادیر خاصی از ماتریس کوواریانس برای انتخاب اینکه کدام یک از خوشه های حاصل نامزد برای تقسیم بعدی است استفاده می کند. این روش تکرار می شود تا زمانی که خوشه مورد نظر به دست می آید. برای خوشه های بزرگ با توزیع گاوسی این یک راه حل بهینه است.

به عنوان مثال، C_1 را به عنوان یک مجموعه ای از پیکسل ها، در نظر بگیرید، تقسیم آن به خوشه های k :

$$(1) \text{ محاسبه } \mu_1, \text{ میانگین } C_1 \text{ و } \Sigma_1, \text{ ماتریس کوواریانس } C_1$$

(2) برای $i = 2$ تا K به صورت زیر عمل کنید:

(i) مجموعه C_n را که دارای بزرگترین مقدار مشخصه است پیدا کنید و بردار مشخصه مرتبط e_n را ذخیره کنید

،

(ii) تقسیم C_n به دو مجموعه، $C_i = \{x \in C_n : e_n^T z_n \leq e_n^T \mu_n\}$ و $C_n^* = C_n - C_i$

(iii) محاسبه $\mu_i, \Sigma_i^*, \mu_n^*$ و μ_i .

5- تقسیم بندی تصویر با استفاده از GrabCut

GrabCut یک گسترش قدرتمند از الگوریتم برش گراف برای تقسیم تصاویر رنگی به طور تکراری است و برای ساده سازی تعامل کاربر برای کیفیت داده شده از نتایج تقسیم بندی مورد نیاز است.



(a)



(b)

شکل 1: مثال تقسیم بندی GrabCut (a) GrabCut اجازه می دهد تا کاربر یک مستطیل را در اطراف جسم

مورد علاقه ی خود که تقسیم شده بود بکشد. (b) جسم تقسیم شده.

GrabCut 5.1 نیمه اتوماتیک اصلی.

الگوریتم GrabCut توزیع های رنگی از پیش زمینه و پس زمینه را با دادن یک احتمال به هر پیکسل که متعلق به خوشه ای از پیکسل های دیگر است یاد می دهد. این را می توان به شرح زیر توضیح داد: یک تصویر رنگی I داده شده است، در نظر میگیریم $Z = (Z_1, \dots, Z_N)$ از N پیکسل، که $Z_i = (C_{1i}, C_{2i}, C_{3i})$ ، $i \in [1, \dots, N]$ و C_j مولفه رنگ j^{th} در فضای رنگ مورد استفاده است. این تقسیم بندی به صورت آرایه $\alpha = (\alpha_1, \dots, \alpha_N)$ ، $\alpha_i \in \{0, 1\}$ تعریف شده است، اختصاص یک برچسب به هر پیکسل تصویر، نشان می دهد که آن متعلق به پس زمینه یا پیش زمینه است. الگوریتم GrabCut عمدتاً شامل دو مرحله اساسی است: اولویت بندی و به حداقل رساندن تکرار. جزئیات هر دو مرحله در بخش زیر توضیح داده شده است.

5.1.1 راه اندازی GrabCut.

تازگی تکنیک GrabCut در "برچسب گذاری ناقص" است که اجازه می دهد یک درجه از میزان تعامل کاربر کاسته شود. این تعامل کاربر به سادگی مشخص کردن فقط پیکسل های پس زمینه با کشیدن یک مستطیل در اطراف شیء پیش زمینه مورد نظر است (شکل 1). فرایند اولیه سازی GrabCut به صورت زیر کار می کند:

مرحله 1: $\text{trimap } A = \{T_B, T_U, T_F\}$ ، در یک روش نیمه اتوماتیک مقداردهی اولیه می شود. این دو منطقه T_B و T_U شامل پس زمینه اولیه و پیکسل های نامشخص است، به ترتیب در حالی که $T_F = \emptyset$. T_B اولیه به عنوان چرخش پیکسل ها در خارج از مستطیل مشخص شده تعیین می شود. پیکسل متعلق به T_B به عنوان یک پس زمینه ثابت در نظر گرفته می شود، در حالی که پیکسل متعلق به T_U توسط الگوریتم برچسب گذاری می شود.

مرحله 2: یک تقسیم بندی تصویر اولیه $\alpha = (\alpha_1, \dots, \alpha_i, \dots, \alpha_N)$ ، $\alpha \in \{0, 1\}$ ایجاد شده، که تمام پیکسل های ناشناخته به طور آزمایشی در کلاس پیش زمینه ($\alpha = 1$ for $i \in T_U$) قرار می گیرند و تمام پیکسل های شناخته شده در این کلاس پس زمینه قرار می گیرند ($\alpha_i = 0$ for $i \in T_B$).

مرحله 3: دو مدل ترکیبی گاوسی کوواریانس کامل (GMMs) تعریف شده است، هر کدام شامل مولفه های K ، یکی برای پیکسل های پس زمینه ($\alpha_i = 0$) و دیگری برای پیکسل های پیش زمینه (در ابتدا ناشناخته) (α_i).

1 (=) است. اجزای K از هر دو GMMs از کلاس های پیش زمینه و پس زمینه با استفاده از تکنیک خوشه بندی درختی و بومان شروع می شوند.

5.1.2. بهینه سازی انرژی تکراری GrabCut.

تقسیم بندی نهایی برش گراف در مراحل زیر با استفاده از الگوریتم بهینه سازی تکراری انجام می شود:

مرحله 4. هر پیکسل در کلاس پیشوند به احتمال زیاد جزء گاوسی در پیش زمینه GMM است. به طور مشابه، هر پیکسل در پس زمینه به احتمال زیاد جز گاوسی پس زمینه اختصاص داده شده است.

مرحله 5. GMM ها دور انداخته می شوند و GMM های جدید نیز در حال یادگیری از مجموعه پیکسل هایی که در مجموعه قبلی ساخته شده اند هستند.

مرحله 6. یک گراف ساخته شده است و گراف Cut برای یافتن یک پیش زمینه جدید و طبقه بندی پس زمینه پیکسل ها اجرا می شود.

مرحله 7. مرحله 4-6 تکرار می شود تا طبقه بندی همگرا باشد.

این مزیت را دارد که امکان پالایش خودکار ابهامات α را فراهم می کند ، به عنوان پیکسل های برچسب گذاری شده جدید از منطقه trimap TU اولیه برای اصلاح رنگ GMM از آن استفاده می شود.

5.2 پیشنهاد GrabCut خودکار.

اگرچه این نقص برچسب گذاری کاربر از GrabCut تعامل کاربر را به طور قابل توجهی کاهش می دهد هنوز برای آغاز روند تقسیم بندی لازم است. این GrabCut را به صورت الگوریتم تقسیم بندی نیمه اتوماتیک / تحت نظارت تعریف می کند. در واقع به تصویر اجازه می دهد که به بخش های مناسب بدون هر گونه راهنمایی کاربر تقسیم شود ، این نیاز به جایگزینی گام نیمه اتوماتیک / تحت نظارت از مقداردهی اولیه GrabCut بصورت کاملاً خودکار / بدون نظارت دارد.

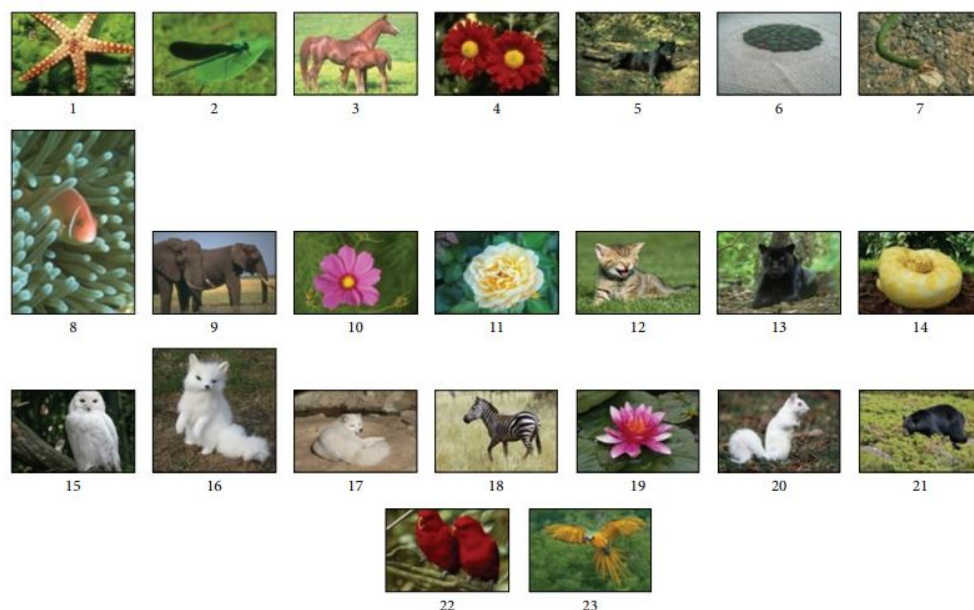
در این مقاله، درختی و بومان پیشنهاد شده است تا به عنوان یک روش خوشه بندی تصویر استفاده شود تا به طور خودکار trimap اولیه T و تقسیم بندی اولیه را تنظیم کند (بخش 5.1، مراحل 1 و 2). تمایز بین trimap و تقسیم بندی رسمی کردن تجزیه منطقه مورد نظر که تقسیم شده است و تقسیم بندی نهایی با الگوریتم GrabCut است. در روش اتوماتیک مراحل 1 و 2 فرآیند مقداردهی اولیه GrabCut به صورت زیر اصلاح خواهد شد.

مرحله 1: در حالی که GrabCut اصلی Trimap T را از دو منطقه TB و TU تشکیل می دهد به عنوان پس زمینه ثابت شده و ناحیه های ناشناخته، به ترتیب، روش اتوماتیک پیشنهادی کل تصویر را به عنوان منطقه ناشناخته TU در نظر می گیرد، که $TU = \{z_i \in \{z_1, \dots, z_n, \dots, z_N\}\}$ و $i \in [1, \dots, N]$. این یعنی که پیش زمینه ثابت یا مناطق پس زمینه شناخته شده نیست و تمام پیکسل های تصویری در کمینه سازی فرایند دخیل خواهند بود تا با الگوریتم برچسب گذاری شود.

مرحله 2: این تصویر با استفاده از تکنیک خوشه بندی درختی و بومان ابتدا به دو پیش زمینه TF و مناطق پس زمینه TB تقسیم می شود. در این مرحله یک GMM جدید معرفی شده است که شامل تنها دو جزء است (K) $= 2$: یک جزء برای پیکسل های پس زمینه ($\alpha_i = 0$) و دیگری برای پیکسل های پیش زمینه ($\alpha_i = 1$). تکنیک خوشه بندی باغچه و Bouman تا رسیدن به تعداد اجزاء ($K = 2$) در GMM اعمال می شود و تکرار می شود، در نتیجه تصویر را کاملاً به دو خوشه تقسیم می کند.

مرحله 3: رنگ پیکسل تصویر متعلق به هر خوشه (خوشه های پیش زمینه و پس زمینه) از مرحله قبلی تولید شده سپس برای مقداردهی اولیه دادن برای دومین بار به صورت کامل به مدل های ترکیبی گواسی کوواریانس (GMMs) با ($K = 5$) استفاده می شود.

مراحل 4-7. بخش یادگیری الگوریتم به صورت GrabCut اصلی دقیقاً اجرا می شود (بخش 5.1، مراحل 4-7).



شکل 2: مجموعه داده ای تصاویر.

6-نتایج و بحث ها

تکنیک GrabCut اتوماتیک به صورت آزمایشی با استفاده از یک مجموعه داده از 23 تصویر مختلف آزمایش شده است، همانطور که در شکل 2 نشان داده شده است. این مجموعه داده ها به طور جزئی از مجموعه داده های تقسیم بندی برکلی و از تصاویر موجود عمومی روشی که مطابق معیارهای خاصی باشد جمع آوری شده اند. این معیارها یک ویژگی خاص را به دو بخش تقسیم بندی طبقه بندی می کند از جمله داشتن عمدتاً یک شی (به عنوان پیش زمینه) و جدایی بین پیش زمینه و مناطق رنگی پس زمینه.

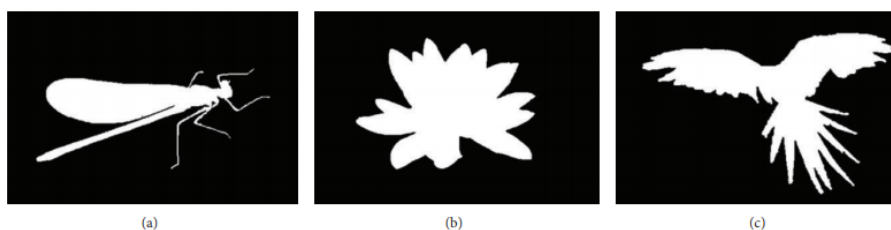
برای ارزیابی، متوجه شدیم که هیچ تقسیم بندی باینری وجود ندارد که به عنوان بخشی از تقسیم بندی های انسانی در مجموعه داده های تقسیم برکلی شامل شود. به این دلیل، داده های حقیقی اساسی برای مجموعه داده های انتخاب شده ما به صورت دستی با استفاده از ابزارهای پردازش تصویر استاندارد (Adobe Photoshop) تولید شده اند. شکل 3 نمونه هایی از داده های دودویی دستی تولید شده را نمایش می دهد. میزان خطا و میزان مره همپوشانی به عنوان دو معیار ارزیابی استفاده می شود. نرخ خطا به عنوان کسری از پیکسل ها با تقسیم بندی اشتباه (در مقایسه با حقایق زمین) تقسیم بر تعداد کل پیکسل ها در تصویر محاسبه شده است. نرخ امتیاز همپوشانی توسط $y_1 \cap y_2 / y_1 \cup y_2$ محاسبه می شود، که y_1 و y_2 هر دو تقسیم بندی باینری است.

در اولین آزمایش، GrabCut اتوماتیک، که مقداردهی اولیه شده است با استفاده از درختی و بومان، با الگوریتم GrabCut اصلی کاربرد و مقایسه می شود. شکل 4 نتایج بصری نمونه را برای تقسیم بندی با استفاده از $K = 5$ GMM ها که توسط Rother و همکاران توصیه می شود نشان می دهد. جدول 1 مقایسه کمی بین GrabCut اصلی و اصلاح شده را برای کل مجموعه داده ها که در شکل 2 ارائه شده است نشان می دهد. همانطور که در جدول 1 نشان داده شده است، GrabCut خودکار با استفاده از خوشه بندی درختی و بومان اصلی از شرایط به حداقل رساندن خطا و بهبود دقت تقسیم بندی را اجرا می کند. نرخ خطای متوسط برای GrabCut اتوماتیک 3.64٪ است نسبت به 4.28٪ برای تکنیک GrabCut اصلی است. عملکرد کلی به نظر می رسد در شرایط انحراف استاندارد (SD) که نشان دهنده 3.61٪ برای GrabCut اتوماتیک در مقایسه با 5.5٪ برای GrabCut اصلی بهتر است.

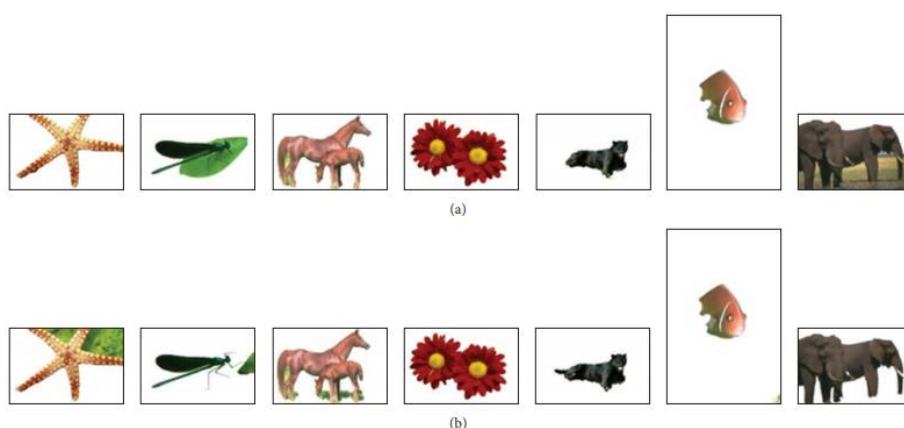
برخی از موارد خطای تقسیم بندی بد با استفاده از GrabCut اصلی را می توان در جدول 1 (تصاویر 1 و 9) مشاهده کرد. این اشکال اصلی از مقداردهی اولیه GrabCut اصلی را توضیح می دهد، که باعث حساسیت نتایج تجزیه و تحلیل به انتخاب کاربر از منطقه موردنظر که باید تقسیم شود می شود. این اتفاق زمانی که دیگر اشیاء، خارج از محدوده مورد نظر که ممکن است به عنوان بخشی از پیش زمینه با قرارگیری درون محدوده مرز مستطیلی کشیده شده در اطراف شیء موردنظر قرار بگیرند می افتد. نتایج تقسیم بندی این دو تصویر در تصویر 4 (a) شرح و نشان داده شده است. می تواند آموزش دهد که چطور یک بخش بزرگی از لایه در تقسیم نهایی تصویر حشرات ظاهر می شود. همین مشکل هنگامی که زمین را به عنوان بخشی از ناحیه پیش زمینه با تصویر فیل در نظر گرفتیم، اتفاق می افتد. مقایسه های کمی میزان خطا تولید شده برای این دو تصویر در جدول 1 و مقایسه های بصری در شکل 4 کارایی GrabCut خودکار در برخورد با چنین مشکل را نشان می دهد. این کارایی GrabCut اتوماتیک با جلوگیری از هرگونه محدودیت سخت که در هنگام مقداردهی اولیه برای هر دو پیش زمینه یا پس زمینه مشخص می شود تحریک می شود (بخش 5.2، مرحله 1).

در آزمایش دوم، GrabCut خودکار، که با استفاده از درختی و بومان، با مدل های فضایی رنگی مختلف مقداردهی اولیه شده بود، از جمله RGB, XYZ, CMY, YUV و HSV اعمال شده بود. ویژگی هایی که هر پیکسل تصویر

را شناسایی می کنند تنها مقداری از این سه جزء در انتخاب فضای رنگی است. نتایج تقسیم بندی نهایی برای همه تصاویر استفاده شده به دست آمده است. برای مقایسه کمی، جدول 2 نرخ خطا و نرخ نمره همپوشانی را برای تمام مجموعه داده ها نشان می دهد. نتایج در جدول 2 به صورت صعودی از چپ به راست در دوره هایی از شماره کلی نتایج تقسیم بندی تصاویر خوب و میانگین نرخ های خطا مرتب شده اند. ما میتوانیم ببینیم که فضای RGB چیزی است که نتایج بهتر را برای بیشتر تصاویر در دوره ها از لحاظ میزان خطای متوسط به دست می آورد. YUV و XYZ را با افزایش بسیار کمی در میزان خطای متوسط دنبال کنید. آنها تقریباً مشابه خطای متوسط و نرخ نمره همپوشانی نشان داده می شوند، که برای نرخ خطا 5.49% و برای نرخ نمره همپوشانی 95.35% است و 5.63% برای میزان خطا و 95.79% برای نرخ امتیاز همپوشانی به ترتیب است. شکل 5 نتایج تقسیم بندی بصری برای برخی از تصاویر را نشان می دهد، در حالی که شکل 6 نمودارهای گراف میانگین میزان خطای تقسیم بندی و نرخ نمره همپوشانی برای همه فضاهای رنگی مختلف را نشان می دهد.



شکل 3: نمونه هایی از داده های زمینی باینری دستی تولید شده.



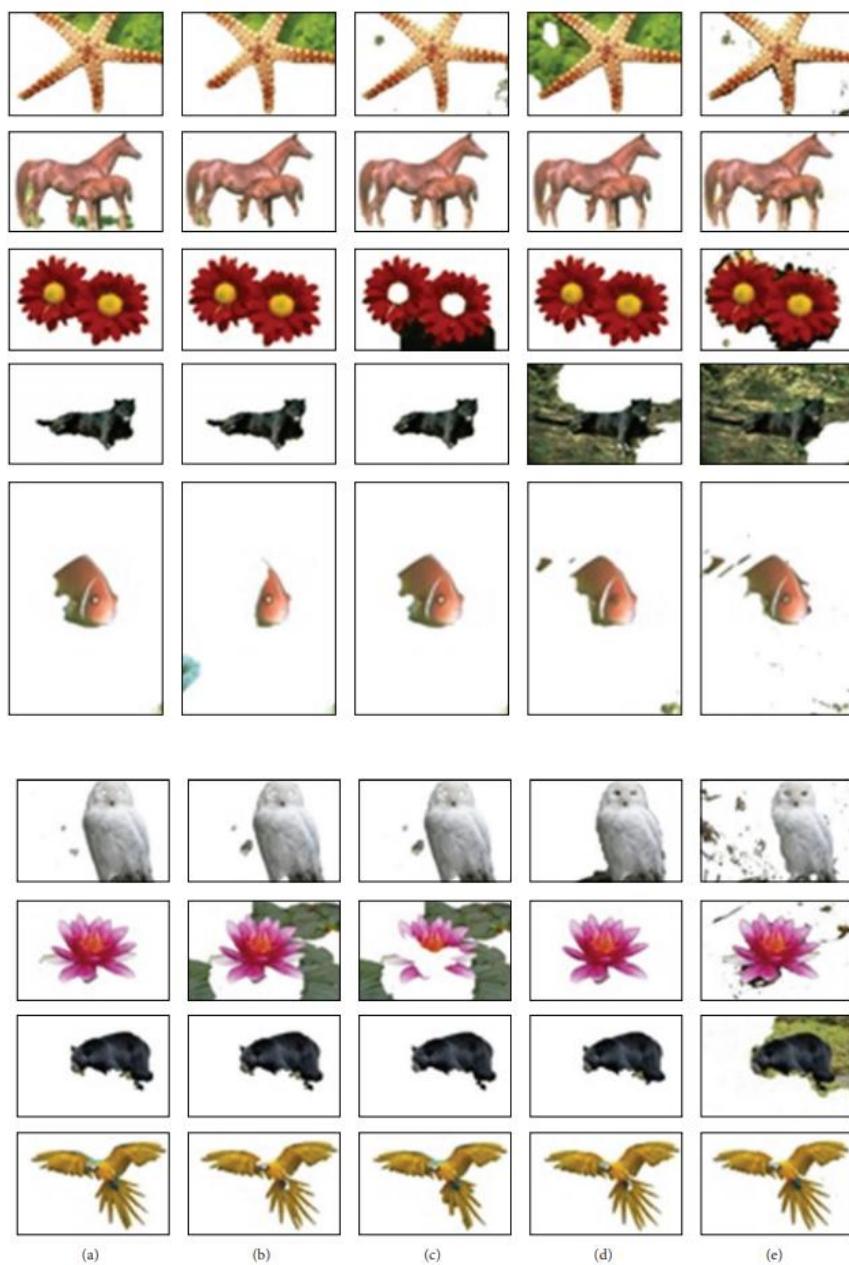
شکل 4: مقایسه بصری نتایج تقسیم بندی از GrabCut (a) نیمه اتوماتیک اصلی و GrabCut (b) اتوماتیک مقداردهی اولیه شده با استفاده از درختی و بومان.

Image	Error rate %		Overlap score rate %	
	Original semiautomatic GrabCut	Automatic GrabCut using Orchard and Bouman	Original semiautomatic GrabCut	Automatic GrabCut using Orchard and Bouman
1	3.05	18.70	95.91	58.57
2	15.48	5.74	43.96	75.69
3	4.79	7.31	91.51	85.48
4	3.07	3.09	97.06	97.02
5	4.16	3.75	82.42	85.18
6	0.86	0.86	97.17	97.17
7	2.40	2.40	69.00	69.01
8	0.87	1.08	92.76	90.81
9	25.81	2.17	67.66	97.31
10	2.82	2.81	96.32	96.35
11	2.05	2.05	97.04	97.04
12	4.99	4.93	88.45	89.32
13	2.28	2.30	95.71	95.64
14	2.36	2.56	96.85	96.41
15	2.78	3.50	94.14	91.98
16	3.16	3.02	93.38	93.80
17	2.08	2.10	95.19	95.11
18	3.88	3.86	90.95	91.06
19	2.88	2.92	93.44	93.30
20	1.43	1.44	96.47	96.43
21	1.27	1.27	94.62	94.64
22	3.30	3.16	93.37	93.73
23	2.57	2.58	93.75	93.74
Avg.	4.28	3.64	89.44	90.21
SD	5.50	3.61	12.76	9.87

جدول 1: مقایسه بین GrabCut اصلی و خودکار.

ترجمه جدول :

نرخ نمره همپوشانی %		نرخ خطا %		تصویر
با استفاده از GrabCut خودکار باغ و بومان	GrabCut نیمه تمام خودکار	با استفاده از GrabCut خودکار باغ و بومان	GrabCut نیمه تمام خودکار	تصویر
				GrabCut نیمه تمام خودکار



شکل 5: مقایسه بصری نتایج تقسیم بندی برای GrabCut اتوماتیک اعمال شده در فضاهاى رنگی RGB (a).

(b) YUV ، (c) XYZ ، (d) CMY و (e) HSV

ترجمه جدول :

نرخ نمره همپوشانی %	نرخ خطا %	تصویر
---------------------	-----------	-------

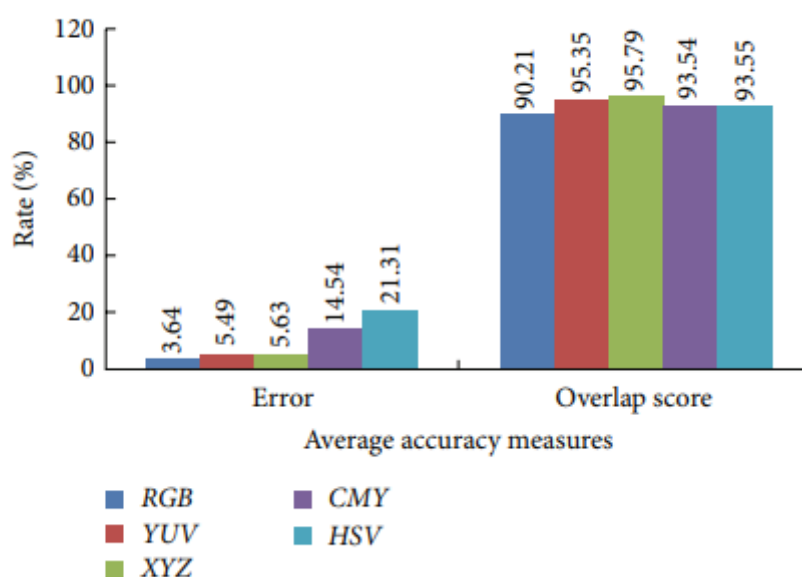
Image	Error rate %					Overlap score rate %				
	RGB	YUV	XYZ	CMY	HSV	RGB	YUV	XYZ	CMY	HSV
1	18.70	20.09	5.45	36.31	5.21	58.57	92.99	98.47	98.05	99.24
2	5.74	2.79	2.92	2.91	19.19	75.69	98.85	98.63	97.95	98.86
3	7.31	5.51	3.90	3.76	5.37	85.48	95.38	99.21	99.31	98.22
4	3.09	3.07	18.95	3.08	12.67	97.02	99.19	88.13	99.11	100
5	3.75	3.76	4.20	42.25	74.38	85.18	89.33	85.70	99.19	96.56
6	0.86	0.86	0.89	30.82	29.90	97.17	99.51	99.56	74.05	98.20
7	2.40	2.33	2.28	1.20	36.11	69.01	99.85	99.76	93.21	98.43
8	1.08	6.94	1.08	2.68	2.68	90.81	44.91	97.27	88.69	88.74
9	2.17	2.16	2.17	31.43	28.80	97.31	99.22	99.24	82.93	85.93
10	2.81	4.19	4.23	4.81	5.69	96.35	99.90	99.92	97.39	100
11	2.05	2.07	2.18	14.99	15.93	97.04	99.91	99.59	63.79	61.16
12	4.93	5.22	4.97	4.44	8.29	89.32	92.48	93.30	96.81	81.42
13	2.30	2.39	2.42	42.31	27.58	95.64	98.73	99.03	99.91	97.09
14	2.56	2.81	3.06	4.15	4.28	96.41	98.37	97.97	94.79	94.55
15	3.50	3.68	3.68	5.27	12.71	91.98	99.28	99.28	99.22	88.40
16	3.02	2.94	2.98	8.78	49.66	93.80	98.89	98.94	99.56	99.79
17	2.10	2.20	2.12	35.89	6.09	95.11	99.07	99.00	99.65	81.17
18	3.86	6.71	6.76	35.85	80.23	91.06	98.95	98.71	96.36	100
19	2.92	37.29	45.78	5.06	6.40	93.30	99.88	64.12	90.29	99.84
20	1.44	1.43	1.47	8.48	27.59	96.43	98.95	99.10	98.04	98.82
21	1.27	1.27	1.04	1.55	22.74	94.64	99.70	98.99	98.18	99.92
22	3.16	3.39	3.39	5.21	5.29	93.73	94.65	94.36	89.97	89.85
23	2.58	3.21	3.56	3.25	3.28	93.74	95.13	94.91	94.95	95.58
Avg.	3.64	5.49	5.63	14.54	21.31	90.21	95.35	95.79	93.54	93.55
SD	3.61	7.92	9.47	15.24	21.64	9.87	11.37	7.84	9.01	9.29

جدول 2: نتایج تقسیم بندی تجربی در فضاهاى رنگى مختلف با استفاده از GrabCut خودکار

7-نتیجه گیری و کار آینده

در این مقاله یک اصلاح از GrabCut برای از بین بردن نیاز به تعامل کاربر اولیه برای هدایت تقسیم بندی و در نتیجه تبدیل GrabCut به یک تکنیک تقسیم بندی اتوماتیک ارائه شده است. این اصلاح شامل استفاده از درختی و بومان به عنوان یک روش خوشه بندی ناپایدار برای مقداردهی اولیه پردازش تقسیم بندی GrabCut است. بر اساس یک مجموعه داده از 23 تصویر، آزمایشات نشان داد که GrabCut خودکار با استفاده از خوشه بندی درختی و بومان از GrabCut اصلی بهتر عمل می کند. این نیاز برای مداخله کاربر را کاهش می دهد درحالی که تقسیم بندی مزیت اضافی برای GrabCut از طریق اتوماسیون را اضافه می کند. علاوه بر این، این تقسیم بندی

دقیق و قوی را با نرخ خطای متوسط از 3.64٪ در مقایسه با نتایج 4.28٪ از نرخ خطای میانگین که توسط GrabCut اصلی بدست می آید فراهم می کند. علاوه بر این، عملکرد از GrabCut اتوماتیک با استفاده از پنج فضای رنگ مختلف، RGB ، XYZ ، HSV ، و CMY ارزیابی می شود. نتایج تجربی نشان می دهد که نتایج تقسیم بندی بسته به فضای رنگی RGB بهترین نتایج تقسیم بندی را در مقایسه با سایر فضاهای رنگی برای مجموعه ای از تصاویر در نظر گرفته شده ارائه می دهد.



شکل 6: مقایسه مقادیر دقت متوسط برای اعمال تقسیم بندی GrabCut خودکار در فضاهای رنگی مختلف.

این مطالعه را می توان با افزایش مجموعه داده ها و از جمله انواع تصاویر مختلف بهبود بخشید. از سوی دیگر، کار آتی ممکن است شامل اصلاح طرز کار به حداقل رساندن انرژی روش GrabCut اتوماتیک که بهینه سازی و تقسیم بندی چند لایه را اجازه می دهد، باشد.

References

- [1] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, Prentice-Hall, 3rd edition, 2006.
- [2] M. Lalitha, M. Kiruthiga, and C. Loganathan, "A survey on image segmentation through clustering algorithm," *International Journal of Science and Research*, vol. 2, no. 2, pp. 348–358, 2013.
- [3] N. Sharma, M. Mishra, and M. Shrivastava, "Colour image segmentation techniques and issues: an approach," *International Journal of Scientific & Technology Research*, vol. 1, no. 4, pp. 9–12, 2012.
- [4] L. Busin, N. Vandenbroucke, and L. Macaire, "Color spaces and image segmentation," *Advances in Imaging and Electron Physics*, vol. 151, pp. 65–168, 2008.

- [5] J. Rother C, V. Kolmogorov, and A. Blake, "GrabCut": interactive foreground extraction using iterated graph cuts," *ACM Transactions on Graphics*, vol. 23, no. 3, pp. 309–314, 2004.
- [6] Y. Y. Boykov and M.-P. Jolly, "Interactive graph cuts for optimal boundary & region segmentation of objects in N-D images," in *Proceedings of the 8th International Conference on Computer Vision (ICCV '01)*, vol. 1, pp. 105–112, IEEE, Vancouver, Canada, July 2001.
- [7] M. T. Orchard and C. A. Bouman, "Color quantization of images," *IEEE Transactions on Signal Processing*, vol. 39, no. 12, pp. 2677–2690, 1991.
- [8] D. Khattab, H. M. Ebied, A. S. Hussien, and M. F. Tolba, "Automatic GrabCut based on unsupervised clustering for image segmentation," *Journal of Computer Science and Technology*. Submitted.
- [9] O. Alata and L. Quintard, "Is there a best color space for color image characterization or representation based on multivariate gaussian mixture model?" *Computer Vision and Image Understanding*, vol. 113, no. 8, pp. 867–877, 2009.
- [10] M. Pagola, R. Ortiz, I. Irigoyen et al., "New method to assess barley nitrogen nutrition status based on image colour analysis: comparison with SPAD-502," *Computers and Electronics in Agriculture*, vol. 65, no. 2, pp. 213–218, 2009.
- [11] A. Jurio, M. Pagola, M. Galar, C. Lopez-Molina, and D. Pater-nain, "A comparison study of different color spaces in clustering based image segmentation," in *Information Processing and Management of Uncertainty in Knowledge-Based Systems. Applications*, vol. 81 of *Communications in Computer and Information Science*, pp. 532–541, Springer, 2010.
- [12] J. M. Chaves-González, M. A. Vega-Rodríguez, J. A. Gómez-Pulido, and J. M. Sánchez-Pérez, "Detecting skin in face recognition systems: a colour spaces study," *Digital Signal Processing*, vol. 20, no. 3, pp. 806–823, 2010.
- [13] C.-J. Du and D.-W. Sun, "Comparison of three methods for classification of pizza topping using different colour space transformations," *Journal of Food Engineering*, vol. 68, no. 3, pp. 277–287, 2005.
- [14] G. Ruiz-Ruiz, J. Gómez-Gil, and L. M. Navas-Gracia, "Testing different color spaces based on hue for the environmentally adaptive segmentation algorithm (EASA)," *Computers and Electronics in Agriculture*, vol. 68, no. 1, pp. 88–96, 2009.
- [15] V. Gulshan, V. Lempitsky, and A. Zisserman, "Humanising GrabCut: Learning to segment humans using the Kinect," in *Proceedings of the IEEE International Conference on Computer Vision Workshops (ICCV Workshops '11)*, pp. 1127–1133, Barcelona, Spain, November 2011.
- [16] A. Hernandez, M. Reyes, S. Escalera, and P. Radeva, "Spatio-Temporal GrabCut human segmentation for face and pose recovery," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW '10)*, San Francisco, Calif, USA, June 2010.
- [17] Y. Hu, *Human Body Region Extraction from Photos*, MVA, 2007.
- [18] D. Corrigan, S. Robinson, and A. Kokaram, "Video matting using motion extended GrabCut," in *Proceedings of the IET European Conference on Visual Media Production (CVMP '08)*, London, UK, 2008.
- [19] C. Görging, B. Fröhlich, and J. Denzler, "Semantic segmentation using GrabCut," in *Proceedings of the International Conference on Computer Vision Theory and Applications (VISAPP '12)*, pp. 597–602, February 2012.
- [20] J. Ramírez, P. Temoche, and R. Carmona, "A volume segmentation approach based on GrabCut," *CLEI Electronic Journal*, vol. 16, no. 2, 2013.
- [21] R. Kaur and G. S. Bhathal, "A survey of clustering techniques," *International Journal of Advanced Research in Computer Science and Software Engineering*, vol. 3, no. 5, pp. 153–157, 2013.
- [22] A. K. Jain, M. N. Murty, and P. J. Flynn, "Data clustering: a review," *ACM Computing Surveys*, vol. 31, no. 3, pp. 316–323, 1999.
- [23] A. Gulhane, P. L. Paikrao, and D. S. Chaudhari, "A review of image data clustering techniques," *International Journal of Soft Computing and Engineering*, vol. 2, no. 1, pp. 212–215, 2012.
- [24] S. Naz, H. Majeed, and H. Irshad, "Image segmentation using fuzzy clustering: a survey," in *Proceedings of the 6th International Conference on Emerging Technologies (ICET '10)*, pp. 181–186, October 2010.
- [25] N. Grira, M. Crucianu, and N. Boujemaa, "Unsupervised and semisupervised clustering: a brief survey," in *Proceedings of the 7th ACM SIGMM International Workshop on Multimedia Information Retrieval*, 2005.
- [26] M. A. Ruzon and C. Tomasi, "Alpha estimation in natural images," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 18–25, IEEE, 2000.
- [27] Y.-Y. Chuang, B. Curless, D. H. Salesin, and R. Szeliski, "A Bayesian approach to digital matting," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '01)*, pp. 11264–11271, December 2001.
- [28] N. Vandenbroucke, L. Macaire, and J.-G. Postaïre, "Color image segmentation by pixel classification in an adapted hybrid color space. Application to soccer image analysis," *Computer Vision and Image Understanding*, vol. 90, no. 2, pp. 190–216, 2003.
- [29] Y. Boykov and V. Kolmogorov, "An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 9, pp. 1124–1137, 2004.
- [30] H. S. Kumar, K. Raja, K. Venugopal, and L. Patnaik, "Automatic image segmentation using wavelets," *International Journal of Computer Science and Network Security*, vol. 9, no. 2, pp. 305–313, 2009.
- [31] C. V. Narayana, E. S. Reddy, and M. S. Prasad, "Automatic image segmentation using ultra fuzziness," *International Journal of Computer Applications*, vol. 49, pp. 6–13, 2012.
- [32] D. Martin, C. Fowlkes, D. Tal, and J. Malik, "A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics," in *Proceedings of the 8th IEEE International Conference on Computer Vision (ICCV '01)*, 2001.
- [33] <http://www.google.com/>.