Kelebek görüntülerin sınıflandırılması için bir içerik bazlı görüntü erişim sistemi

Yılmaz KAYA[a1]; Lokman KAYCİ[b]

[a]*Bilgisayar Mühendisliği Bölümü, Siirt Üniversitesi, 56100 Siirt*

e-posta: yilmazkaya1977@gmail.com

[b]*Biyoloji Bölümü, Siirt Üniversitesi, 56100 Siirt*

e-posta: lokmankayci@mynet.com

Özet

Kelebekler kendi aralarında sınıflandırılırken öncelikle kelebeklere ait kanat, baş, anten gibi dış morfolojik özelliklerine göre sınıflandırılır. Dış morfolojik özelliklerine göre sınıflandırmanın gerçekleştirilemediği durumlarda teşhis için genital özelliklerin incelenmesi gerekir. Kelebeklerin genital yapılarının incelenmesi farklı kimyasal maddeler ve yöntemler uygulanarak sağlanabilir. Bu işlemler ancak belli maliyetlerle gerçekleştirilebilir. Ayrıca genital preparatların hazırlanması belli aşamalar gerektirdiği için zaman alıcıdır. Bu çalışmada kelebek teşhisi için klasik teşhis yöntemlerine alternatif web tabanlı bir içerik tabanlı görüntü erişim (İTGE) görü sistemi önerilmiştir. Çalışmada *Papilinidae, Pieridae* ailelerinin *38* kelebek türüne ait 380 görüntü kullanılmıştır. Kelebeklerin karşılaştırılması için yapısal benzerlik indeks metodu (structural similarity index measurement (SSIM)) kullanıldı. Elde edilen sonuçlara göre önerilen yöntemin kelebek görüntülerin sınıflandırılmasında önemli başarı gösterdiği saptanmıştır.

Anahtar Kelimeler: *kelebek sınıflandırma, yapısal benzerlik ölçütü, bilgisayarla görü, örüntü tanıma*

[1] Baş yazar. E-Posta: yilmazkaya1977@gmail.com (Yılmaz Kaya).

1-Giriş

Kelebek ve güveler takımı, sahip olduğu 170.000 den fazla türüyle böcekler arasındaki en zengin takımlardan biridir. Türkiye, 5000'den çok *Lepidopter* türü ile bölgenin zengin ülkeleri arasında sayılır [1]. Kelebeklerde kanat şekilleri, desen ve renkler son derece büyük farklılıklar gösterir. Öyle ki, bu özellikler ilk bakışta çoğu defa türlerin ayrımında önemli rol oynarlar. Türler içinde sınırlı kaldığı sürece bu tip özellikler taksonomik karakterler olarak kullanılırken, bazen birbirine dıştan çok benzeyen türlerin ayrımında, bilhassa bireyin dış genital organlarının yapısal özelliklerinin incelenmesine ihtiyaç duyulur [1].

Bilgi miktarındaki hızlı artışa paralel olarak bilgi erişimle ilgili sorunlar da artmakta ve çeşitlenmektedir. Büyük miktarlardaki bilgi yığınları içinde kullanıcıların istenilen bilgiye erişmesi zorlaşmaktadır [2].

Günümüzde metinsel erişimle ilgili problemleri bile tam anlamıyla çözülmemişken, karşımıza görsel bilgiye erişimdeki sorunlar çıkmaktadır. Sayısal görüntü miktarındaki artış ve kullanıcıların deneyimlerindeki sınırlamalar, görsel bilgi erişim sistemlerinin sorunlarını artırmaktadır. Son yıllarda WWW üzerinden bilgiye erişilmekte. Bilgi yığınları arasında doğru bilgiye ulaşılmaya çalışılmaktadır. Görsel bilgiler ile görüntülerin sorgulanması içerik tabanlı görüntü erişimi (İTGE, content-based image retrieval =CBIR) olarak isimlendirilir [3]. İTGE, bir görüntünün büyük görüntü yığınları içerisinde bilgisayar görü teknolojileri ile sorgulanmasıdır.

İTGE sistemler, iki görüntü arasındaki renk, desen, doku gibi özellikleri karşılaştırarak görüntüleri bir benzerlik ölçütü ile ayırmaktadır[4].

Bu çalışmanı amacı, kelebek kanatlarındaki doku, desen yapılarını kullanarak kelebekleri otomatik sınıflandıran bir bilgisayar görü sistemi geliştirmektir. Desen, bir görüntüde piksel ton uzaysal değişimini belirten ve farklı alanlarda sayısal görüntülerin sınıflandırılmasında kullanılabilen yapılardır [5]. Kelebek kanat desenleri farklı türlerde farklı renk, parlaklık, desende olabilir. Dolayısıyla bilgisayar görüsü yöntemlerinde kanat desenleri matematiksel olarak ifade edilebilirler. Benzerlik yöntemleri dokusal özellikler işleyen yöntemlere göre daha basit, hızlı ve insan görme sistemine (Human Vision System=HVS) daha yakın olduklarından [6] dolayı kelebeklerin karşılaştırılması için yapısal benzerlik indeksi (YBİ, Structural Similarity Index Measurement, SSIM) kullanılmıştır. YBİ, iki görüntüye ait parlaklık, kontrast, ve yapı özelliklerini karşılaştırarak iki görüntü arasında bir benzerlik indeksi oluşturmaktadır [7]. YBİ, Wang ve arkadaşları tarafından önerilmiş olup bir çok problemde başarılı bir şekilde kullanılmıştır [8]. Çalışmada 38 farklı türe ait 380 kelebek

görüntüsü kullanılmıştır. Elde edilen sonuçlara göre geliştirilen İTGE sisteminin kelebek türünün tespitinde önemli başarı sağladığı görülmüştür.

2. Veri Seti

Çalışma ile ilgili inceleme materyali Van, Erek Dağı'nda toplanmıştır. Kelebekler arazide tül atrap kullanılarak yakalanmıştır. Etil asetat içeren kavanozlarda öldürüldükten sonra önceden hazırlanmış özel zarflarının içerisine gerekli toplama bilgilerini içeren etiketiyle birlikte konmuştur. Örnekler standart müze materyali şeklinde hazırlanmak üzere germe tahtalarında gerilerek kurutulmuştur. Dış morfolojik özelliklere bağlı olarak yapılan teşhislerde, baş ve göğüs üzerindeki organ ve ekstremitelerin morfolojik özelliklerinin yanı sıra kanatların alt ve üst yüzlerindeki renk ve desenler dikkate alınmıştır. Türlerin teşhisinde çeşitli el kitapları, revizyon çalışmaları ve karşılaştırma materyalleri kullanılmıştır. Teşhiste kullanılan yayınlar alfabetik sıraya göre şu yazarlara aittir: [9,10,11,12]. Çalışmada kullanılan türlerin doku tanıma yöntemine göre teşhis edilmesi için her türe ait digital Nikon profesyonel makine ile çekilmiş yüksek çözünürlüklü 10'ar fotoğraf kullanılmıştır. Bu çalışmada *Papilinidae, Pieridae* kelebek ailelerinin Şekil 1'de gösterilen 38 türü kullanılmıştır.



Şekil 1: Örnek kelebek görüntüleri

3. Yapısal Benzerlik Indeksi Ölçütü (Structural Similarity Index Measurement)

Bu yöntem, iki görüntüye ait parlaklık, kontrast ve yapı denilen 3 özniteliğin karşılaştırılmasına dayanmaktadır [8]. Bu yöntemde parlaklık (luminance) (L), kontrast (contrast) (c), ve yapı (structure) (s) özellikleri birbirinden bağımsız olarak elde edilmektedir. *x* ve *y* birer görüntü olmak üzere YBİ ölçütü aşağıdaki eşitlik ile elde edilmektedir.

$$SSIM(x,y) = f(l(x,y),c(x,y),s(x,y)) \quad (1)$$

burada $l(x,y)$ parlaklık karşılaştırma fonksiyonu olup eşitlik 2 ile elde edilmektedir.[8];

$$l(x,y) = \frac{2\mu_x\mu_y + c_1}{\mu_x^2 + \mu_y^2 + c_1} \quad (2)$$

burada, $c_1 = (K_1 L)^2$, K_1 küçük bir sabit değer ($K_1 << 1$) ve L ise piksel dağılım ölçütüdür.

$c(x,y)$, kontrast karşılaştırma fonksiyonu:

$$c(x,y) = \frac{2\sigma_x\sigma_y + c_2}{\sigma_x^2 + \sigma_y^2 + c_2}, \quad (3)$$

burada, $c_2 = (K_2 L)^2$, K_2 küçük bir sabit değer ($K_2 << 1$)ve L ise piksel dağılım ölçütüdür.

$s(x,y)$, yapı karşılaştırma fonksiyonu;

$$s(x,y) = \frac{\sigma_{xy} + c_3}{\sigma_x\sigma_y + c_3} \quad (4)$$

Son olarak YBİ ölçütü aşağıdaki eşitlik ile elde edilmektedir [13]:

$$YBİ(x,y) = [l(x,y)^{\alpha}.c(x,y)^{\beta}.s(x,y)^{\gamma}] \quad (5)$$

burada, $\alpha > 0, \beta > 0$ *ve* $\gamma > 0$

Eğer $\alpha = \beta = \gamma = 1$ ve $c_3 = c_2/2$, ise;

$$YBİ(x,y)\frac{(2\mu_x\mu_y + c_1)(2\sigma_x\sigma_y + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad (6)$$

dir. Burada x,y karşılaştırılan görüntüler, μ_x, σ_x ve σ_{xy} tahmin edilen görüntüye ait piksel yoğunluk ortalaması, standart sapmasını ve ortak varyansı belirtmektedir.

$$\mu_x = \frac{1}{N}\sum_{i=1}^{N} x_i \tag{7}$$

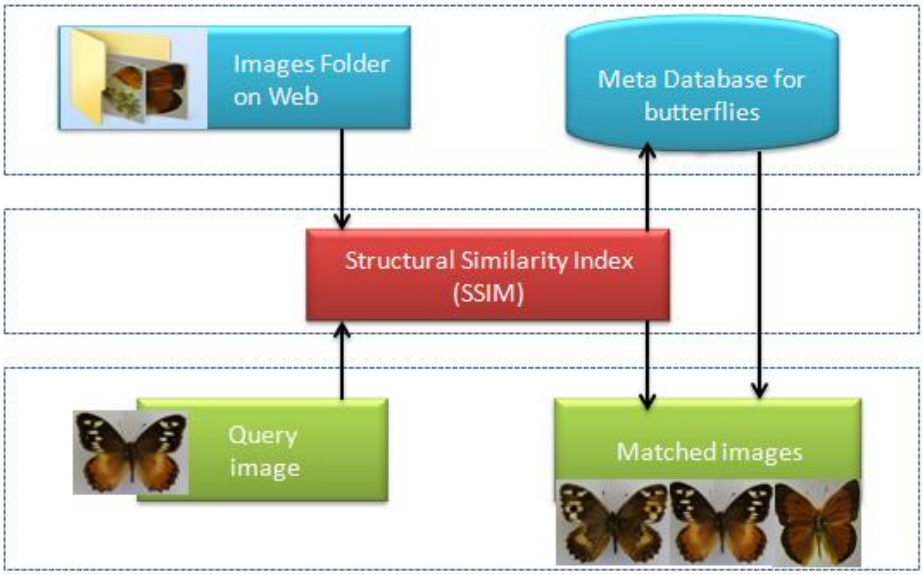
$$\sigma_x = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(x_i - \mu_x)^2} \tag{8}$$

ve

$$\sigma_{xy} = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \mu_x)(y_i - \mu_{y)} \tag{9}$$

4. Geliştirilen ITGE Sistemi

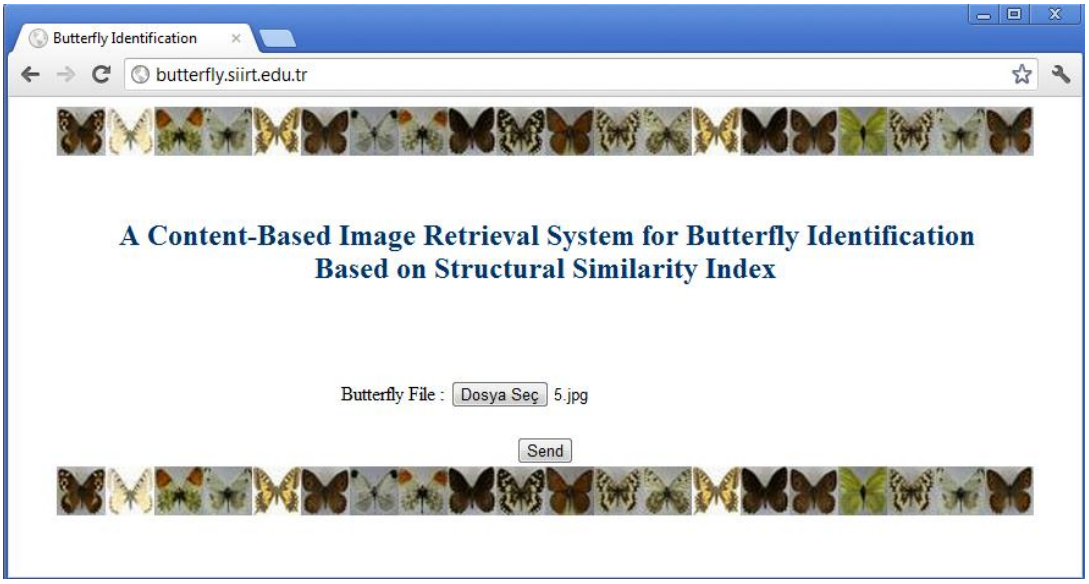
Bu çalışmada geliştirilecek İTGE sistemine ait diyagram şekil 2'te gösterilmiştir. Sistemin çalışma aşamaları aşağıdaki basamaklarda özetlenmiştir.



Şekil 2: Kelebek teşhisi için İTGE sistemi

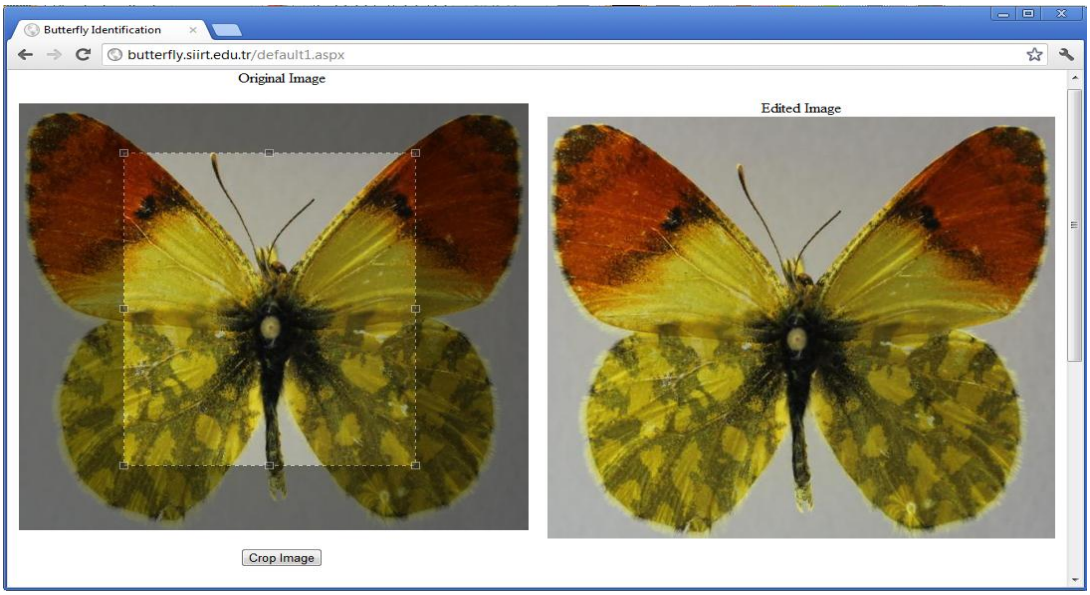
Kelebek teşhisi için geliştirilen ITGE sistem aşamaları:

1) İnternet kullanıcısı web üzerinden geliştirilen şekil 2'de gösterilen web sayfasına giriş yapar.



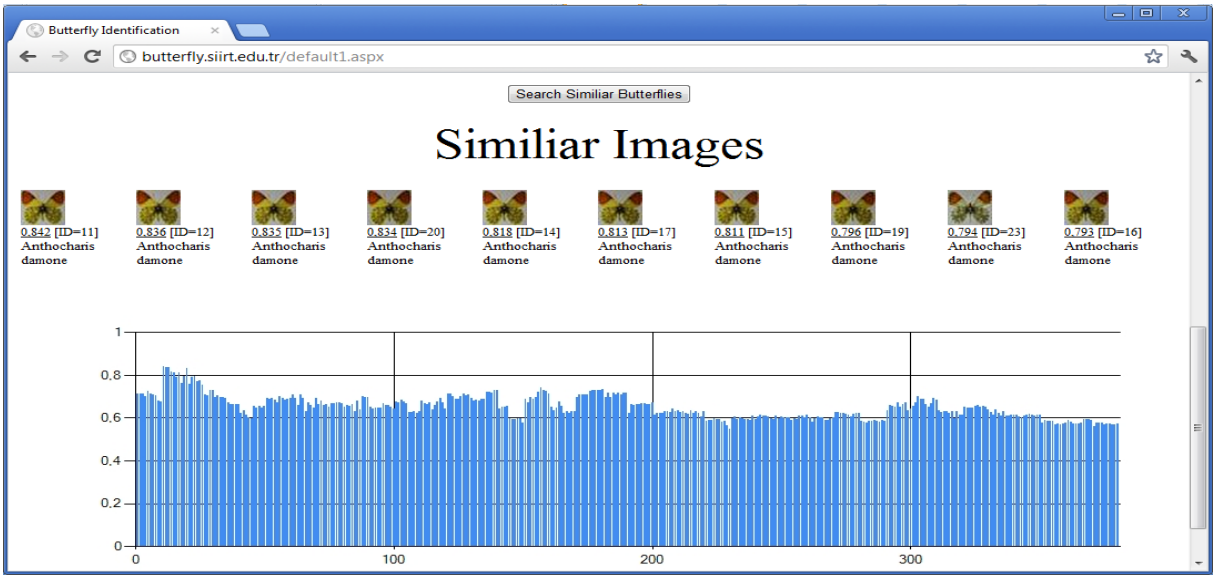
Şekil 3: Teşhis edilecek kelebek görüntüsünün seçimi.

2) Kullanıcı teşhis edilecek görüntüyü web sayfasına yükledikten sonra ilgili alanın belirlenmesi için kırpma işlemi gerçekleştirilir (Şekil 4). Kırpma işlemi, görüntü üzerindeki gereksiz alanların atılması için performans açısından önem arz etmektedir.



Şekil 4: İlgilenilen alanın belirleme aşamasına örnek.

3) Kelebek görüntüsünden gereksiz alanlar atıldıktan sonra "Benzer Kelebek Görüntülerini Ara" düğmesi tıklanarak arama işlemi başlatılır. Geliştirilen İTGE sistemi sorgulanan görüntüsünün benzediği en iyi 10 kelebek görüntüsünü göstermektedir (Şekil 5).

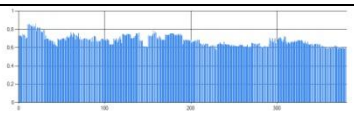
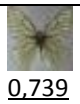
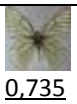
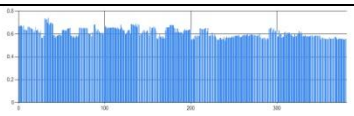
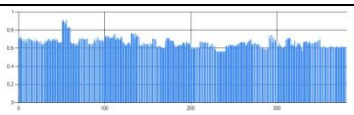


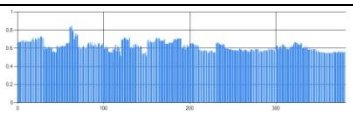
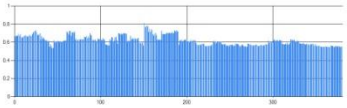
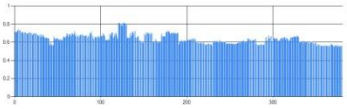
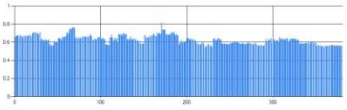
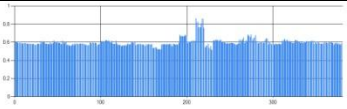
Şekil 5: Sorgu görüntüsüne benzeyen görüntülerin listesi.

5. Sonuçlar

Bu çalışmada yapısal benzerlik metodu kullanılarak bir ITGE sistemi geliştirilmiştir. Yapısal benzerlik metodu iki görüntünün parlaklık, kontrast ve yapı bileşenlerine bağlıdır. Çalışmada 38 farklı kelebek türüne ait 380 görüntü kullanıldı. Rastgele seçilen görüntülerin benzerlik gösterdiği görüntülerin listesi Tablo 1'de verilmiştir. Benzerlik ölçütü değerleri histogramlar şeklinde gösterilmektedir.

Tablo 1: Rastgele seçilen kelebek görüntülerin benzerlik gösterdiği görüntüler

Deney No	Sorgulanan Görüntü	Benzer görüntü ve SSIM değeri					Benzerlik Histogramı
E1	 Anthocharis damone	 0,861	 0,858	 0,853	 0,852	 0,849	
E2	 Aporia crataegi	 0,739	 0,735	 0,727	 0,725	 0,698	
E3		 0,917	 0,914	 0,896	 0,896	 0,887	

	Colias crocea						
E4	 Euchloe penia	 0,849	 0,835	 0,826	 0,80	 0,730 *	
E5	 Pontia callidice	 0,812	 0,788	 0,777	 0,773	 0,768 *	
E6	 Pieris krueperi	 0,818	 0,812	 0,811	 0,808	 0,807	
E7	 Pontia edusa	 0,814	 0,769 *	 0,760 *	 0,738	 0,738	
E8	 Papilio machaon	 0,861	 0,861	 0,857	 0,826	 0,817	
E9	 Chazara bischoffi	 0,702	 0,701	 0,691	 0,550 *	0,553 *	
E10	Pseudochazara beroe	0,655	0,652	0,552 *	0,548 *	0,539 *	

*Yanlış sınıflandırılan görüntüler

10 adet rastgele seçilen kelebek görüntüleri için denemeler gerçekleştirildi. Tablo 1'de görüldüğü gibi sorgulanan görüntülerin veri tabanında bulunan görüntüler ile benzerlik değerleri 0,655 ve 0,917 arasında değişmiştir. Histogramlara bakıldığında sorgulanan kelebek görüntülerin benzerlik gösterdiği görüntüler arasında benzerlik ölçütünün yüksek

çıktığı, benzerlik göstermedikleri görüntüler ile aralarında benzerlik ölçütünün düşük çıktığı görülmektedir. E4, E5, E7, E9, ve E10 deneyleri için yanlış kelebekler de sorgulanmıştır. Ancak tüm deneylerde ilk listelenen kelebekler doğru görüntüler olduğu görülmektedir.

6. Tartışma

Görüntü bazlı arama uygulamaları günümüzde önem kazanan hizmetler olmaktadır. Günümüzde iki yaklaşım söz konusu, birincisi görüntüye ait metinsel meta bilgilere dayanan yöntemler iken, diğer görüntü içeriğini değerlendiren yöntemlerdir. Görüntülere ait görsel özelliklerin kullanılarak arama işlemleri gerçekleştiren sistemlere içerik tabanlı görüntü erişimi (İTGE) sistemler olarak isimlendirilmektedir. Bu çalışmada, kelebek teşhisi için YBİ benzerlik yöntemi kullanılarak bir İTGE sistemi geliştirildi. Kelebek görüntülerin kanat desenleri, renkleri ilk etapta kelebek teşhisi için kullanılmaktadır. Ancak her zaman bu yöntem başarılı olamamaktadır. Bu yüzden klasik yöntemlere alternatif yöntemlere ihtiyaç duyulmaktadır. Literatüre bakıldığında bilgisayar destekli yöntemlerin yeterince kullanılmadığı görülmektedir. Bu çalışmada otomatik tanıma işlemi gerçekleştiren bir bilgisayar görü sistemi önerilmiştir. Çalışmada yapılan 10 farklı deneyde sonuçların başarılı olduğu saptanmıştır. Sonuç olarak YBİ yöntemi kullanılarak geliştirilen İTGE sistemi kelebek görüntülerin teşhisinde önemli başarı sağladığı ve benzer uygulamalar için kullanılabileceği öngörülmektedir.

Teşekkür

Bu çalışma, Siirt Üniversitesi Bilimsel araştırma Projeleri Başkanlığı (SÜBAPB) tarafından desteklenmiştir. Proje no: 2013-SİÜMÜH-M1.

References

[1]. L. Kayci, "Erek Dağı (Van) Papilionoidea ve Hesperioidea Ekolojisi ve Faunası Üzerine Araştırmalar (Lepidoptera)", Priamus Supplement, Vol. 6,pp. 1-47, 2007.

[2]. X. R. Marcela, H.B. Pedro, J.T. Caetano, M.A. Paulo, A.R. Natalia, J.M. Traina, "Supporting content-based image retrieval and computer-aided diagnosis systems with association rule-based techniques", Data & Knowledge Engineering, Vol. 68, pp. 1370–1382, 2009.

[3]. K. Yuan, Z. Tian, J. Zou, Y. Bai, Q. You, “Brain CT image database building for computer-aided diagnosis using content-based image retrieval”, Information Processing and Management, Vol. 47, pp. 176–185, 2011.

[4]. Y. Liu, D. Zhang, G. Lu, W. Ma, „A survey of content-based image retrieval with high-level semantics”, Pattern Recognition, Vol. 40, pp. 262 – 282, 2007.

[5]. W. Y. Ma, B.S. Manjunath, “A texture thesaurus for browsing large aerial photographs”, Journal of the American Society for Information Science, Vol. **49**(7), pp. 633-648, 1998.

[6]. J. R. Bach, “A visual information management system for the interactive retrieval of faces”, IEEE Transactions on Knowledge and Data Engineering**,** Vol**. 5**(4), pp. 619-628, 1999.

[7]. L. Chaofeng, C. B. Alan, “Content-partitioned structural similarity index for image quality assessment”, Signal Processing,Image Communication, Vol. 25, pp. 517–526, 2010.

[8]. Z. Wang, A.C. Bovik, H.R. Sheikh, E.P. Simoncelli, “Image Quality Assessment: From Error Visibility to Structural Similarity”, IEEE Transactions On Image Processing, Vol. 13(4), pp. 600-612, 2004.

[9]. F. Carbonell, “C ontribution a la connaissance du genre Agrodiaetus Hübner (1822), position taxinomique d'Agrodiaetus anticarmon Koçak, 1983 (Lepidoptera, Lycaenidae) ” Linneana Belgica, Vol. 16 (7),pp. 263-265, 1998.

[10]. H. Paul, G. Ryan, “The Promise of DNA Barcoding for Taxonomy”, Systmtc. Biolgy, Vol. 54 (5), pp. 852-859, 2005.

[11]. P.Skala, “New taxa of the genus Hyponephele MUSCHAMP, 1915 from Iran and Turkey (Lepidoptera, Nymphalidae)”, Linneana Belgica, Vol. 19 (1), pp. 41-50, 2003.

[12]. T. Tolman, “Butterflies of Britain and Europe. Harper Collins Publishers, London. 320 (1997).

[13]. Z. Wang, A.C. Bovik, “A universal image quality index”, IEEE Signal Processing Letters, Vol. 9, pp. 81–84, 2002.