Research Article/Araştırma Makalesi

gazi JOURNAL OF ENGINEERING SCIENCES

Gazi AKADEMİK YAYINCILIK

GAZİ
JOURNAL OF ENGINEERING SCIENCES

Secure Message Transmission Via Object Detection Verification Using Images Encoded With Lehmer Algorithm Based Random Key

Simay Hoşmeyve*[a], Arda Cem Bilecan[b], Bahadır Karasulu[c], İsmet Ünlü[d]

Submitted: 28.10.2022 Revised: 11.02.2023 Accepted: 04.03.2023 doi:10.30855/gmbd.0705057

ABSTRACT

Nowadays, the successful transmission of data between end users, data protection and confidentiality are important issues. In our study, a system that ensures the transmission of bitwise encrypted images in pixel color matrix information based on the Lehmer algorithm for the generation of random keys over an end-to-end reliable communication channel and guarantees accuracy validation has been created. Using the deep learning model, metadata is created with the properties of the detected objects in the image scene. In the study, structured data is created in Base64 encoding format after encryption is done by the sender end of the message using the Lehmer algorithm. This message is transmitted end-to-end to the receiving party over the secure transmission channel. Finally, the encoded version of the image and the metadata information are extracted from the structured data exchange and transfer file containing the detection information of the objects in the image scene in the message packet reaching the receiver end, and also, decryption is performed. On the receiver end, object detection results of the deep learning model are compared with the decrypted information obtained from this file, and an objective verification of the accuracy of the message is also performed. Considering the average values in the experiments, the average values with the data obtained from three different data sets are for the Number of Pixels Change Rate (NPCR) 99.61%, the Unified Average Changing Intensity (UACI) 14.70%, and also, the maximum average entropy value is obtained 7.9999 for encrypted images. In our study, discussions and evaluations based on scientific findings are included.

Keywords: Random number, encryption, deep learning, data communication

[a,*] EyeCU Vision,
20816-Maryland,
Amerika Birleşik Devletleri.
Orcid: 0000-0001-8478-3126
e mail: simayhosmeyve@gmail.com

[b] NTT Data Business Solutions,
34746 – Ataşehir, İstanbul, Türkiye.
Orcid: 0000-0002-0455-337X

[c] Çanakkale Onsekiz Mart University,
Engineering Faculty,
Dept. of Computer Engineering.
17020 – Çanakkale, Türkiye.
Orcid: 0000-0001-8524-874X

[d] Smartin Information
Technologies LTD,
34306 – Başakşehir,
İstanbul, Türkiye.
Orcid: 0000-0002-6949-8666

*Corresponding author:
simayhosmeyve@gmail.com

Lehmer Algoritması Tabanlı Rastgele Anahtar ile Şifrelenmiş Görüntüleri Kullanarak Nesne Tespitli Doğrulama Yoluyla Güvenli Mesaj İletimi

ÖZ

Günümüzde son kullanıcılar arası verinin başarılı iletimi, verinin korunması ve gizlilik önemli konulardır. Çalışmamızda, uçtan uca (end-to-end) güvenilir bir iletişim kanalı üzerinden rastgele anahtar üretimi için Lehmer algoritması tabanlı piksel renk matris bilgilerinde bit bazında şifrelenmiş görüntülerin iletilmesini sağlayan ve doğruluk geçerlemesini garanti eden bir sistem oluşturulmuştur. Derin öğrenme modeli kullanılarak görüntü sahnesi içerisindeki tespit edilen nesnelerin özellikleri ile üst veri oluşturulmaktadır. Çalışmada, mesajın göndericisi tarafından Lehmer algoritması kullanılarak şifreleme yapıldıktan sonra Base64 kodlama formatında yapılandırılmış veri oluşturulmaktadır. Güvenli iletim kanalı üzerinden bu mesaj alıcı tarafa uçtan uca iletir. Son olarak alıcı tarafa ulaşan mesaj paketindeki görüntü sahnesi içerisindeki nesnelerin tespit bilgilerini içeren yapılandırılmış veri değişim ve aktarım dosyasından görüntünün kodlanmış hali ve üst veri bilgileri ayıklanarak şifre çözümlemesi yapılmaktadır. Alıcı tarafta derin öğrenme modelinin nesne tespit sonuçları bu dosyadan elde edilen de-şifrelenmiş bilgi ile kıyaslanarak mesajın doğruluğunun geçerlemesi şeklinde nesnel sağlaması da yapılmaktadır. Yapılan deneylerdeki ortalama değerler göz önüne alındığında üç farklı veri kümesinden elde edilen veriler için ortalama değerler olarak Piksel Sayısı Değişim Hızı (NPCR) %99,61, Birleşik Ortalama Değişme Yoğunluğu (UACI) %14,70 ve ortalama entropi değeri ise şifrelenmiş görüntüler için azami 7,9999 değerinde ölçülmüştür. Çalışmamızda bilimsel bulgulara dayanan tartışma ve değerlendirmelere yer verilmektedir.

Anahtar Kelimeler: Rastgele sayı, şifreleme, derin öğrenme, veri iletişimi

To cite this article: S. Hoşmeyve, A. C. Bilecan, B. Karasulu and İ. Ünlü, "Secure Message Transmission Via Object Detection Verification Using Images Encoded With Lehmer Algorithm Based Random Key," *Gazi Journal of Engineering Sciences,* vol. 9, no. 1, pp. 108-127, 2023. doi:10.30855/gmbd.0705057

1. Giriş (Introduction)

Günümüzde veri miktarının artmasıyla beraber çeşitli ortamlardan (ses, görüntü, video) elde edilen anlamlandırılması gereken ham veriler bulunmaktadır. Bu anlamlandırma için ilk önce var olan ses veya video sahnesindeki nesnelerin sağlıklı bir biçimde tespit edilmesi gerekmektedir. Ardından veri güvenliği için veri iletimine temel oluşturması adına şifreleme mekanizmaları ve algoritmaları sıklıkla kullanılmaktadır.

Çalışmamızda veri şifrelemesini sağlamak için sözde rastgele sayı üretici olarak Lehmer Algoritması seçilmiştir [1]. Üretilen sözde rastgele sayılar ile üretilen anahtarlar görüntülerin şifrelenmesinde kaynak olarak kullanılmıştır. Üretilen anahtarların mantıksal Özel VEYA (eXclusive OR, XOR) işlemlerinde kullanılabilmesi için görüntünün her bir piksel değeri göz önüne alınarak 8 bit değerine uygun normalizasyon işlemi de yapılmıştır. Böylece Lehmer algoritması tabanlı elde edilen rastgele anahtar görüntü şifreleme ve deşifrelemede başarıyla kullanılmıştır.

Mesaj Kuyruklama Telemetri İletimi (Message Queueing Telemetry Transport, MQTT) [2, 3] protokolü, iletişime dahil olan makinalar arasında (machine-to-machine, M2M) mesaj gönderimi tabanlı güvenli bir iletişim protokolüdür. Daha çok Nesnelerin Interneti (Internet of Things, IoT) [4] ekosistemlerinde kullanılır. Nesnelerin Interneti fiziksel nesneler arası veya dış sistemlerle oluşturulan veri alışverişi veya nesneler arası senkronizasyon sağlayan bir iletişim ağıdır. Bu protokol, İstek ile Yanıt yapısına [5] dayalı olmakla beraber abone için yayın (broadcast-to-subscriber) bağlantısında TCP/IP [6, 7] bağlantısı kurar. Bu sayede pek çok işletim sistemlerinde kullanılabilir.

Derin öğrenme, düşük seviyeli özniteliklerin elde edilmesiyle oluşturulan belirli bir sınıflandırma esnasında daha yüksek seviyeden anlamsal olarak zengin özniteliklerin öğrenilmesini amaçlayan günümüzde literatürdeki çalışmalarda sıklıkla kullanılan başarımı yüksek bir makine öğrenmesi alt alanıdır. Bu alt alandaki modeller, yoğunlukla yapay sinir ağlarının katmanlı yapısının uygun sayıda hesapsal öğe ve katman ile derinleştirilerek, verinin filtrelenmesi ve belirgin özniteliklere dayanan daha doğru bir sınıflandırma ve tespit işleminin yapılabilmesini sağlamaktadır. Buna literatürde sıklıkla *Temsili Öğrenme (representation learning)* denilmektedir [8].

Sadece Bir Kez Bak (You Only Look Once, YOLO) derin öğrenme modeli, nesne tespiti ve sınıflandırma alanında son yıllarda oldukça popülerliği artması nedeniyle hem etkinliği hem de sınıflandırma başarımının yüksek oluşu ndan dolayı literatürdeki çalışmalarda tercih edilir olmuştur [9]. YOLO modeli transfer öğrenmeyi de kullanması nedeniyle önceden eğitimli olarak çok sayıdaki sınıfı görüntüdeki sahne içerisindeki kalan nesneleri sınırlayıcı kutu tabanlı olarak tespit ve sınıflandırmasını yapabilmektedir. Çalışmamızda YOLO derin öğrenme modeli kullanılarak deneylerde kullanılan görüntülerin sahne içeriklerindeki nesneler hakkındaki bilgiler şifrelemesi yapılan görüntüler ile birlikte alıcı makinelere üst veri (metadata) biçiminde JavaScript Nesne Gösterimi (JavaScript Object Notation, JSON) formatlı paket içerisinde MQTT güvenli iletişim kanalı üzerinden iletilmektedir.

Zıt Dil Görüntü Ön Eğitimi (Contrastive Language-Image Pretraining, CLIP) derin öğrenme modeli ile çalışmamızda önceden eğitilmiş derin sinir ağır modellerinin ağırlıkları kullanarak sıfır örnek (zero shot) ile öğrenme yöntemi ile kullanıcının gönderebileceği görüntü çeşitliliğine karşı başarılı bir şekilde tahmin edebilmek hedeflenmiştir [10]. CLIP modeli girdi olarak görüntü verisini alarak metin ile eşleştirmekte ve etiket olarak görüntüde en yoğun olarak bulunan sınıfı metin olarak vermektedir. MQTT bağlantısı ile veri iletiminde gönderilecek mesaj içerisinde görüntü üzerinden elde edilen metinsel sonuç alıcı için gönderinin başarılı bir şekilde alındığına dair bir doğrulama aracı olarak kullanılmaktadır.

Çalışmamızın ana amacı uçtan uca (end-to-end) güvenilir bir iletişim kanalı üzerinden Lehmer algoritması tabanlı şifrelenmiş görüntülerin iletilmesini sağlayan bir sistem oluşturulmasıdır. Görüntü iletimine ve görüntülerdeki nesne çeşitliliğine dayanan doğrulama ve geçerleme başarımı tüm sistem için göz önüne alınmıştır. Bu nedenle sistemin iletilen verideki şifreleme ve deşifreleme doğruluğu çeşitli ölçütlerle ölçülmüştür. Literatürde, şifre kırmaya yönelik diferansiyel ve lineer saldırıya dayanıklılık başarımı olarak görüntü şifreleme yönteminin başarısını ölçen bahsi geçen ölçütler ile sistemin ana amacına ulaştığı deneylerle kanıtlanmıştır. Ayrıca, ilgili nesne sınıfları bazında nesne tespitine dayanan başarımı çalışmamızdaki deneysel sonuçlarla ortaya konulmuştur. Literatürdeki

çalışmalara bakıldığında, sadece bit bazında (bitwise) mantıksal Özel VEYA (eXclusive OR, XOR) işlemi tabanlı şifreleme yaklaşımları, sözde rasgele sayı üretiminde Doğrusal Eşleşmeli Üreteçler (Linear Congruential Generator, LCG) kullanımı, YOLO [12] derin öğrenme modeli kullanılarak sadece ilgilenilen bölge (Region of Interest, ROI) üzerinden kısıtlı bilgiyle şifreleme yapılan deneyler bulunmaktadır [11, 13-15]. Bu açıdan önceki çalışmaların aksine sadece ilgilenilen bölge değil, çalışmamızdaki deneylerde kullanılan tam ölçekli görüntülerin YOLO derin öğrenme modeli kullanılarak sahne içeriklerindeki nesneler hakkındaki bilgilerin tam olarak çeşitlerine göre elde edilmesiyle oluşan üst veri (metadata) kullanılmıştır. Bu üst verinin yanı sıra bu şifrelemesi yapılan görüntüler de Base64 ile kodlanarak (Base64 encoding) şifrelenmiş görüntünün iletildiği alıcı makinelere JavaScript Nesne Gösterimi (JavaScript Object Notation, JSON) formatlı veri paketi içerisinde güvenli iletişim kanalı üzerinden iletilmektedir. Çalışmamızın ele aldığı güvenli iletim kanalı üzerinden şifrelenmiş tam ölçekli görüntülerin iletimi ve deşifreleme ile doğrulaması adına nesne tespitinde derin öğrenmenin kullanımı yoluyla uygun bir sistemin oluşturulması sayesinde literatürdeki çalışmalara karşın temel farklılık ve öne çıkan katkı oluşmaktadır. Çalışmamızın literatüre ana katkısı; genel kullanıcı için güvenli kanaldan iletilen şifrelenmiş görüntüdeki nesnelerin sınıflarının doğru tespitine dayanan bir şifreleme ilâ deşifreleme tabanlı nesne tespiti geçerlilemesini garanti etmesidir. Buna göre; literatürde sıklıkla kullanılan başarım oranı yüksek derin öğrenme modeline dayanan altyapıyla kullanıcının görüntü iletiminde nesne çeşiti kısıtlılığına bağlı kalmadan güvenilir iletim kanalı üzerinden şifre kalitesi yüksek olduğu ilgili saldırıya dayanıklılık ölçütleriyle betimlenmiştir. Bu sayede görüntüleri mesajlaşma yoluyla ileten bu güvenilir sistemi kullanıcının sorunsuzca kullanımının sağlanmıştır.

Bu çalışma altı bölümden oluşmaktadır. İkinci bölümde literatürdeki önceki çalışmalara yer verilmekte, üçüncü bölümde materyal ve yöntem olarak Lehmer algoritması ile rastgele sayı üretimi, derin öğrenme ile nesne tespiti için derin sinir ağı modellerinden bahsedilmektedir. Dördüncü bölümde çalışmamızda temel altyapıyı oluşturan önerilen sistemin detaylarına yer verilmektedir. Beşinci bölümde deneysel sonuçlara yer verilirken, altıncı bölümde ise bilimsel olgulara dayanan tartışma ve değerlendirme sunulmaktadır.

2. Önceki Çalışmalar (Related Work)

Sözde rastgele sayı üreteçleri ile görüntü şifreleme üzerine literatürde yapılan çalışmalar incelendiğinde Banthia ve Tiwari [11] 2013 yılındaki çalışmasında Doğrusal Eşleşmeli Üreteçler (Linear Congruential Generator, LCG) kapsamında "*ImageEncryptLCG*" adıyla geçen "*Doğrusal Uyumlu Üreteç*" algoritması ve "*ImageEncryptionChaos*" adıyla geçen kaotik haritalama algoritmaları bazlı olan algoritma ile orijinal şifrelenmiş görüntüler arasında bilimsel ölçütler karşılaştırmasını incelemişlerdir. Banthia ve Tiwari ismi geçen iki yöntemin de seçilen doğru parametre değerleri ile başarılı sonuçlar verdiğini gözlemlemişlerdir. Çalışmamızda kullanılan Lehmer üreteci, "ImageEncryptLCG" algoritmasının düzenlenmiş bir çeşididir.

Viswanatha ve arkadaşları [12] 2022 yılındaki çalışmasında YOLOv4'ün mimarisinden, performans ölçütlerinden ve veri kümesinin içeriğinin başarım sonuçlarına etkisinden bahsetmişlerdir. Viswanatha ve arkadaşlarının hayvan sınıflandırması üzerinde yaptıkları testte veri kümesinin büyütülmesi ile daha özelleştirilmiş ayrımların yapılabildiğinden ve nesne konumlandırmalarının sonuçlarda etkili olduğunu tespit etmişlerdir. Araştırmacılar yaptıkları versiyon karşılaştırmalarında YOLO'nun ikinci ana sürümü olarak bilinen YOLOv4'ün hız ve doğruluk için en yüksek sonuçları verdiği sonucuna varmışlardır.

Oğraş ve Tür [13] 2022 yılındaki çalışmasında etkili bir görüntü şifrelemesi yapmak amacıyla yeni yöntemler oluşturup denemişlerdir. Bu çalışmada kaotik haritalar olarak bilinen "Lojistik Harita" ve "Sinüs Haritası" temel alınarak yeni bir kaotik harita oluşturulmuştur. Oğraş ve Tür'ün oluşturdukları kaotik harita iyi bir şifreleme için gerekli olan rastgeleliği oluşturmayı başarmışlardır. Çalışmada bit tersine çevirme yöntemi ve yeni kaotik harita kullanılarak daha fazla karmaşıklığa sahip ve öngörülemez şifrelemeler yapılmıştır.

Somaraj ve Hussain [14] 2014 yılındaki çalışmasında görüntü şifreleme ve deşifreleme için XOR işleminin iki yöntemle kullanmışlardır. Çalışmada tıbbi görüntülerin şifrelenmesi amaçlanmıştır. İlk yöntemde görüntü XOR işlemi kullanılarak anahtar görüntü ile şifrelenmiştir. İkinci yöntemde anahtar görüntüsünün bit düzlemlerinden biri orijinal görüntüde şifreleme için kullanılmış ve karıştırma

yapılmıştır. Kullanılan bu yöntemler ölçüm sonuçlarında başarılı kabul edilmiştir ve farklı görüntü türlerinde de uygulanabilirdir. Kang ve Choi [15] 2021 yılındaki çalışmasında görüntü şifreleme sisteminin hesaplama maliyetini azaltmayı ve yeterli güvenlik sağlamayı amaçlamışlardır. Görüntüde önemli görülen bölge YOLO modeli ile tespit edilmiş ve böylece şifreleme yapılacak ilgi alanı belirlenmiştir. YOLO modeli ile tespit edilerek kırpılan bölge kaotik bazlı ve yüksek güvenlikli bir sistem ile şifrelenmiştir. Kang ve Choi şifreleme hızını arttırmak için görüntüde piksel karıştırma işlemi uygulamışlardır.

Tüm bu çalışmalar göz önüne alındığında çalışmamızdaki oluşturulan sistem, uçtan uca (end-to-end) güvenilir bir iletişim kanalı üzerinden Lehmer algoritması tabanlı şifrelenmiş görüntüleri ve bunların yanı sıra ilgili sahne hakkındaki üst veriyi (metadata) içerecek şekilde iletilir. Bu sistem son kullanıcı tarafından kullanımında tekrar edilebilir ve sürdürülebilir güvenlik isteklerini karşılayabilecek bir bilişim sistemi altyapısı için tercih edilebilir olmaktadır.

3. Materyal ve Metod (Material and Method)

Bu bölümde çalışmamızdaki alt yapıyı oluştururken kullanılan metodlar, şifreleme, derin öğrenme modelleri ve güvenli iletişim kanalının oluşturulmasındaki protokolün detaylarına yer verilmektedir.

3.1. Lehmer algoritması ile rastgele sayı üretimi (Random number generation with lehmer algorithm)

Literatürde şifreleme alanında birçok çalışma vardır. Özellikle, "Sözde Rastgele Sayı Üreteçleri" (Pseudo Random Number Generator, PRNG) [16] rastgele sayı dizilerine yaklaşan sayı dizileri üreten üreteçlerdir. Gerçek rastgele sayı üretimi için ise dış bir kaynaktan veri almak gereklidir çünkü dış kaynaklar determinist değildir ve bilgi teorisinden bilinen olasılıksal karışıklık ölçütü olarak entropi daha fazladır. Sıcaklık, fare ve klavye hareketleri, saat bu üreteçler için kaynak (tohum) olabilir. Sözde Rastgele Sayı Üreteçleri gerçek hayattaki rastgeleliğe ve entropi değerlerine ürettiği sayı dizilerinde yaklaşmayı hedefler. Bir başlangıç değeri yani tohum (seed) seçilerek daha sonra bu tohuma tekrarlı olarak uygulanan algoritma ile yeni rastgele sayılar üretilebilir.

Lehmer Algoritması [17], diğer donanımsal ve yazılımsal yaklaşımlara nazaran oldukça az hesapsal karmaşıklık ile yazılımsal sonuç üreten Sözde Rastgele Sayı Üreteçleri'nden bir tanesidir. Bu algoritmada kullanılan formül özyinelemeli olarak devam etmekte böylece istenen sayıda rastgele sayı üretilmeye devam etmektedir. Algoritmanın sonucu olarak gerçek rastgele sayıların rastgeleliğine en yakın hale getirmeye çalıştığımız rastgele sayı dizisi üretilmektedir. Bu dizi daha sonra görüntünün şifrelenmesi için gerekli anahtar üretiminde kullanılmaktadır. Şifrelenmemiş haldeki ham görüntünün matris biçimindeki bilgisinin bit bazında bu anahtara uygun olarak yüksek karıştırma oranıyla şifrelenmesi çalışmamızda hedeflenmiştir. Lehmer algoritması, Doğrusal Eşleşmeli Üreteçler (Linear Congruential Generator, LCG) [18] kategorisinde yer alır. Doğrusal Eşleşmeli Üreteçler'de üretilen rastgele sayısal değer ile sonraki konuma aynı algoritma kullanılarak yeni sayısal değerin üretilmesi sağlanır. Bahsedilen üreteçlerde kullanılan genel formül aşağıda Denklem (1)'de gösterildiği gibidir.

$$x_k + 1 = (ax_k + c) \times mod(m) \quad (1)$$

Lehmer algoritmasında Denklem (1)'deki ifadeye göre ise c sabiti sıfır değeri alınır. Bu nedenle Lehmer algoritmasının genel formülü aşağıda Denklem (2)'de gösterildiği gibidir.

$$x_k + 1 = (ax_k) \times mod(m) \quad (2)$$

Denklem (2)'ye göre m terimine verilecek sayı asal veya asal bir sayının üssü olmalıdır. Buradaki a terimi ise sabit bir çarpandır ve yine asal sayı olarak seçilir. Lehmer algoritmasına göre x_0 terimi genel formüldeki x_k terimine dayanarak tanımlaması yapılan ve $0<x_0< m$ aralığında olması gereken bir başlangıç değeridir. Bu terim değerleri incelenirken, m terimine göre modu alındığında kalansız bölünmemesi için m ile x_0 aralarında asal olmasına dikkat edilir. Buradaki, x_0 terimi tohum (seed) olarak kullanılır. Bahsedilen Lehmer algoritmasındaki değişkenler çalışmamızda sırasıyla; m için 215, c için sıfır değeri olurken, 2 ilâ 19997 aralığından seçilerek formüle uygun olarak a ve x_0 terimleri için asal sayılar kullanılmıştır. Çalışmamızda kullanılan her bir görüntü üç kanal renkli olarak 256×256×3 boyutlarındadır. Buna göre deneylerimizdeki veri kümelerinden alınan her bir görüntü toplam 196608 pikselden oluşacak şekilde bu çözünürlüğe ayarlanmıştır. Her piksel ise 0 ilâ 255 sayıları arasında

değerler almakta olup, bir piksel değeri 8 bit olarak belirlenmiştir. Her bir piksel için Lehmer Algoritması ile Sözde Rastgele Sayı oluşturulmuştur. Bunun ardından bu sözde rastgele sayıların normalizasyon işlemi yapılır. Böylece tüm değerler 0 ilâ 255 aralığına normalize edilmiş olmaktadır. Sözde rastgele sayı kullanılarak oluşturulan değerlerle orijinal görüntüden alınan 196608 adet pikselin mevcut renk değerleri üzerinden bit bazında (bitwise) mantıksal Özel VEYA (eXclusive OR, XOR) işlemi yapılmaktadır. Bu sayede XOR işleminin ardından şifrelenmiş görüntü, orijinal görüntünün mevcut bilgisi karıştırılmış hale getirilerek elde edilmiş olmaktadır.

3.2. Derin öğrenme ile nesne tespiti (Object detection with deep learning)

Derin öğrenme altyapısının oluşturulmasında temel olarak evrişimli sinir ağı modelleri, metin ve görüntü kodlayıcı (encoder) temelli derin sinir ağı modelleri kullanılmıştır. Bununla ilgili detaylara aşağıda yer verilmektedir.

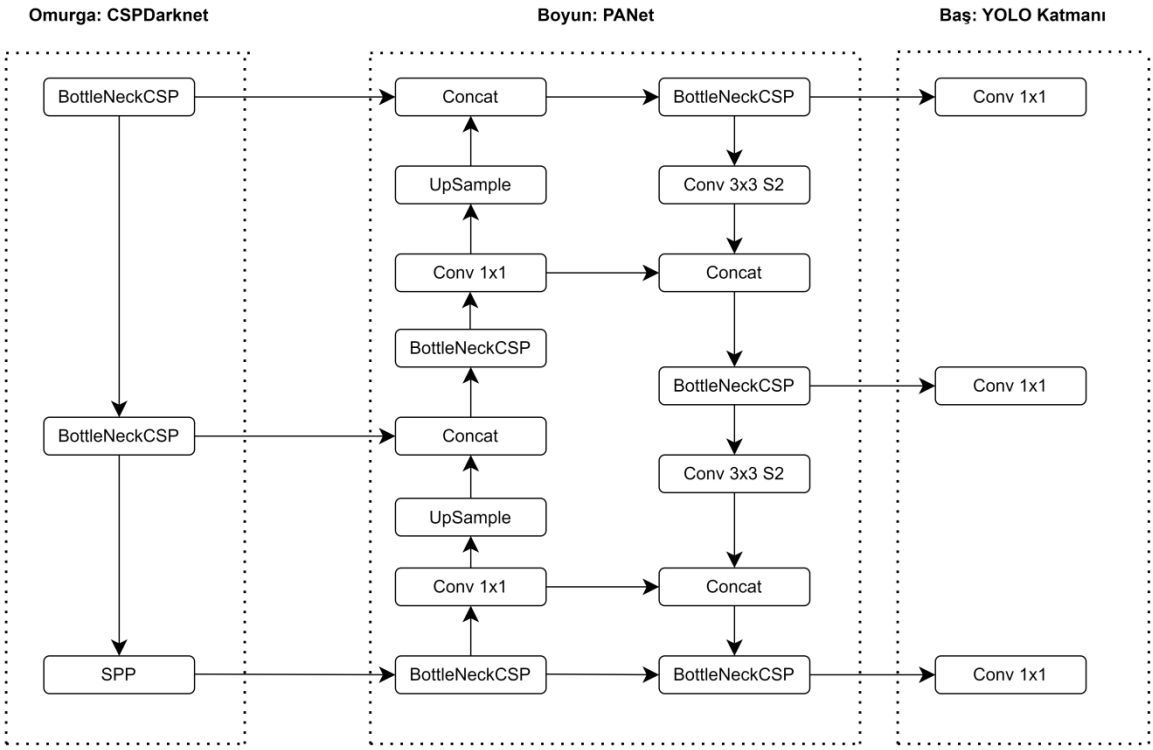
3.2.1. Sadece bir kez bak modeli (You only look once model)

Sadece Bir Kez Bak (You Only Look Once, YOLO) derin sinir ağı modeli gerçek zamanlı nesne tespitinde kullanılan bir derin öğrenme modelidir [9, 12, 15, 19]. Nesne tespiti işlemini, belirli nesnelerin bulunduğu görüntüdeki konumunu belirlemenin yanı sıra bu nesneleri sınıflandırmayı da içerir. Daha önceki çalışmalarda (Artıklı Evrişimli Sinir Ağı, Residual Convolutional Neural Network, R-CNN ve varyasyonları), bu işlem birden fazla adımda bir işlem hattı üzerinden yapılmaktadır, fakat her bir bileşenin birbirinden ayrı eğitilmesi gerekliliği, çalıştırılma süresinin uzunluğu ve eniyilenmesindeki bazı zorluklar gibi kısıtlayıcı unsurlar ortaya çıkmaktadır. YOLO modelinde bu işlemler tek bir sinir ağında gerçekleşmektedir.

Çalışmamızda YOLO modelinin bir versiyonu olan YOLOv5 kullanılmıştır. Çalışmamızda YOLOv5 modelinin ana katkısı gözetilerek sahne içerisindeki nesnelerin tespitinde kullanılmıştır. Bunun sebebi model boyutunun esnek kontrolü, özel bir aktivasyon işlevinin uygulanması ve veri iyileştirmesi yapabilmesidir [9]. Bu modelin bu versiyonu üç temel bileşenden oluşmaktadır. Bunlar; Baş (head), boyun (neck) ve omurga (backbone) olarak bilinir. YOLOv5 modelinde baş bileşeni YOLO katmanına, boyun bileşeni Yol Kümeleme Ağı'na (Path Aggregation Network, PANet), omurga bileşeni Çapraz Aşamalar Arası Kısmi Bağlantılar'a (Cross-Stage-Partial-Connections, CSPDarknet) denk gelmektedir. Omurga bileşeni içerisinde Uzamsal Piramit Biriktirme (Spatial Pyramid Pooling, SPP) alt bileşeni bulunmaktadır. Böylece bu bileşen verilen görüntünün temel özniteliklerini tespit ederek ve bunları işlenmesi adına evrişimli katmanlardan oluşan ağın bir parçasıdır. DarboğazCSP (BottleneckCSP) alt bileşeni çeşitli evrişim katmanlarından oluşarak model derinliğini arttırarak öznitelik kalitesini daha da geliştirmektedir.

Omurga bileşeni ilk önce ImageNet [20-22] gibi bir sınıflandırma veri kümesi ile eğitilir ve algılama, sınıflandırmadan daha ince ayrıntılar gerektirdiğinden, genellikle son algılama modelinden daha düşük olan bir çözünürlükte eğitilmektedir. Boyun bileşeni içerisinde gerekli aşamalarda Birleştirme Fonksiyonu (Concatenate Function) kullanılmaktadır. Bu boyun bileşeni, olasılıklar ve sınırlayıcı kutu (bounding box) koordinatlarıyla ilgili tahminler yapmak adına tamamen bağlantılı katmanlara sahip omurgadaki evrişim katmanlarından gelen öznitellikleri kullanır.

Baş bileşeni, transfer öğrenimi için giriş katmanı ile aynı şekil (shape) değerlerine sahiptir, diğer katmanlarla değiştirilebilen ağın son çıkış katmanıdır. Baş için temel formül olarak $S \times S \times (K + (B * 5))$ formülüne sahip olan bir tensör verilmektedir. Orijinal YOLO makalesinde [9, 15, 19] belirtildiği gibi, bu tensörün boyutları $7 \times 7 \times 30$ şeklinde bir matristir. Bölünmüşlük boyutu S terimi ile ifade edilmekte ve 7 olarak alınmaktadır. Sınıfı ifade eden K terimi ise 20 değerini almakta ve öngörülen sınırlayıcı kutu terimi ise B olarak temsil edilmekte ve 2 değerini almaktadır. Anlaşılacağı üzere modelin temel üç bileşeni, önce görüntüden temel görsel öznitelikleri elde etme, ardından sınıflandırma işlemi yapma ve sahne içerisindeki ilişkileri ortaya koyma adına birlikte çalışmaktadır. Aşağıdaki Şekil 1 'de YOLOv5 modeline ait derin sinir ağı mimarisi şematik olarak verilmektedir.



Şekil 1. YOLOv5 modeline ait derin sinir ağı mimarisinin blok diyagramı [19] (Block diagram of the YOLOv5 model's deep neural network architecture)

3.2.2. Zıt dil görüntü ön eğitimi modeli (Contrastive language–image pre-training model)

Zıt Dil Görüntü Ön Eğitimi (Contrastive language–image pre-training, CLIP) derin sinir ağı modeli [23, 24, 25], doğal dil denetimi ile ilgili görüntüleri eşleştiren ve öğrenen bir yapay sinir ağıdır. Metin ve görüntü eşleşmesini tahmin etmek için bir görüntü kodlayıcı ve metin kodlayıcı önceden eğitilir. Bu eğitim sıfır örnekli (zero-shot) bir sınıflandırıcı oluşturmak için kullanılır. Sıfır örnekli özelliği sayesinde önceden eğitilmiş 400 milyon görüntü ve metin çifti veri kümesinin eğitimi ile herhangi özelleştirilmiş bir model eğitimi yapmadan tahmin üretebilmeyi sağlar. Deneylerimizde, mevcut sahne için öğrenme sonucunda elde edilen temel yorumlamaya dayalı etiketleme yapmak için CLIP modeli kullanılarak yüksek doğrulukta nesne sınıfları tahminlenmiş, böylece bu bilgi bir üst veri olarak son kullanıcıya aktarılmıştır.

4. Önerilen Sistem (Proposed System)

Bu bölümde şifreleme ve veri iletimi altyapısını oluşturan sistemin hangi aşamaları olduğu ve bu aşamaların teknik detaylarının, uygulama esaslarının neler olduğu açıklanmaktadır.

4.1. Veri haberleşmesi (Data communication)

Literatürde veri iletimi ile ilgili birçok protokol bulunmaktadır. Mesaj Kuyruklama Telemetri İletimi (Message Queueing Telemetry Transport, MQTT) protokolü [2, 3] makinalar arası olan (M2M) mesaj tabanlı bir iletişim protokolü olup daha çok akıllı ev (smart home) sistemleri gibi sistemlerde kullanılan cihazların birbirleriyle güvenli ve hızlı bir şekilde iletişim kurulmasını sağlamaktadır. Bu iletişim yönteminde en önemli aktör MQTT Aracılar'dır (Broker). Aracılar'ın asıl görevleri ise Aboneler'e mesajlar göndermektir. Yayıncı'dan (Server) mesajlar alınır ve Aracılar'a verilir. Aracılar ise gelen mesajı hangi İstemci'ye (Client) göndermesi gerektiğini mesaj içeriğinde bulunan Konu'ya (Topic) bakar. Aboneler'in ise abone oldukları Konu veya Konular vardır. Aracılar gelen mesajın Konu'suna bakarak, o konuya abone olmuş İstemciler'e mesajı yollamaktadır. Çalışmamızda, MQTT protokolü ile görüntülerin uçtan uca ve akranlar arası (end-to-end and peer level) iletiminin daha da güvenli hale getirebilmek adına, "*Lehmer Rastgele Sayı Üretim Algoritması*" ile üretilen anahtarlar sayesinde görüntüler şifrelenip ilgili istemcilere belirli bir JSON paketi formatında gönderilmiştir.

4.2. Video iletimi (Video transmission)

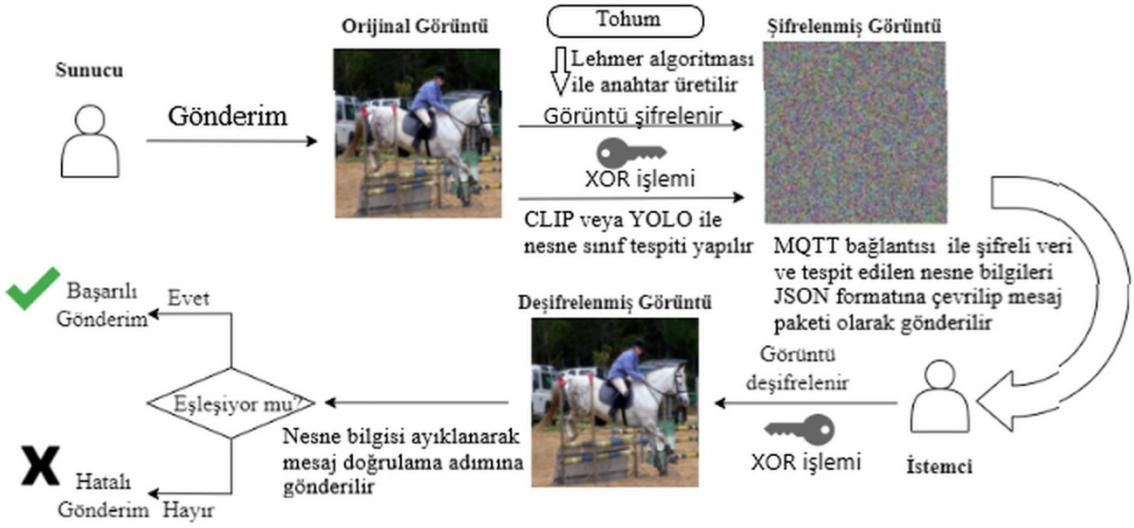
Çalışmamızda birçok çerçeveden oluşan videoyu tek seferde iletilmesinin iletim maliyetini ve işlem karmaşıklığını artıracak olmasından dolayı videodan anahtar çerçeve elde edip daha sonrasında bu görüntüleri birleştirerek bir taşıyıcı ana çerçeve (container frame) elde edilmesi tercih edilmiş, böylece bunun iletim işleminin yapılması sağlanmıştır. Anahtar çerçevelerin bir video parçasından edinimi işlemi Python kütüphanesi olan Katna [26] modülü ile gerçekleştirilmiştir.

Katna modülü videoyu, *CIE* 1976 *L*, u*, v** renk uzayında mutlak farklılıkları, elde edilen video çerçevelerinin parlaklık puanı, elde edilen çerçevelerin entropi, kontrast puanı, görüntü histogramları bilgilerini kullanarak işleme tabi tutmaktadır. Buna göre, video çerçevelerinin *k*-ortalama kümelemesi; Laplace işleci (görüntü bulanıklığı ile ilgili işlemler için) ve değişkenliğine göre kümeler arasından en iyi çerçevenin seçilmesi gibi kriterlere ve işlemlere tabi tutarak videodaki anahtar çerçevelerin elde edilmesini sağlamaktadır.

Çalışmamızdaki deneylerde kullanılan videolar üzerinde Katna kullanılarak her bir video için belirleyici unsurlar içeren 16 anahtar çerçeve elde edilir, bu çerçeveler Python Görüntüleme Kütüphanesi (Python Imaging Library, PIL) modülü [27] kullanımı yoluyla 4×4'lük bir matris formunda görselleştirilen çözünürlüğü 1280×960 olan tek bir görüntünün oluşturulmasıyla iletim işlemine hazır hale getirilmektedir. Bu taşıyıcı ana çerçeve diğer tek çerçeve (durağan görüntüler) için olan işlemlerde de olduğu gibi Lehmer algoritması tabanlı şifrelemeyle ve MQTT bazlı güvenli iletişimde kullanılmaktadır. Çalışmamızda deneyler bu taşıyıcı ana çerçeveler ile gerçekleştirilmiştir.

4.2. Sistem mimarisi (System architecture)

Çalışmamızda videolardan elde edilmiş çeşitli anahtar çerçevelerden oluşan bazı çoklu görüntü içeren taşıyıcı ana çerçeve görüntüler ve farklı görüntü veri kümelerinden elde edilen tek çerçeve görüntüler kullanılarak nesne tespiti yapılmıştır. Tek çerçeve görüntüler için CLIP derin öğrenme modeli, videolardan elde edilen taşıyıcı ana çerçeve görüntüler için ise YOLO derin öğrenme modeli kullanılmıştır. Çalışmamızda geliştirilen sistem üç ana kısımdan oluşmaktadır. Bu kısımlar sırasıyla şifreleme-deşifreleme, veri iletimi ve nesne tespitidir. Buna dair tümleşik sistemin blok diyagramı Şekil 2 ile verilmektedir.



Şekil 2. Önerilen sistemin mimarisinin blok diyagram (Block diagram of the proposed system's architecture)

Şekil 2'den görülebileceği gibi deneylerdeki görüntüler 3 kanal renkli girdi görüntüsü şeklinde verildikten sonra görüntü matris halinde renk uzayındaki her bir kanal kullanılarak oluşturulan matris bir boyutlu vektöre dönüştürülerek ilgili bilgiyi oluşturan ikili (binary) dizi permütasyon yoluyla karıştırılarak kullanılır. Bu karıştırılmış görüntünün üzerine tohum (seed) yoluyla beslenen Lehmer algoritması ile oluşturulan hesapsal duyarlılığı 64 bit kayar nokta (floating point) *double duyarlık* tipinde (veri temsili için kullanılan Python programlama diliyle IEEE-*754* standardına uygun biçimde en önemli 53 bit temsiline indirgenerek kullanılmıştır) [17] bir sözde rastgele sayı tabanlı anahtar

kullanılarak görüntüdeki üç ayrı renk kanalının her birinin üzerine kanal bazında bit seviyesinde mantıksal Özel VEYA (eXclusive OR, XOR) işlemi yapılmaktadır [28]. Bahsi geçen IEEE-*754* standardına uygun tip indirgemesine bağımlı olarak üretilen anahtarın değer uzayı (key space) oluşmaktadır. Bu yolla, orijinal görüntünün şifrelenmiş hali bu anahtarla oluşturulmakta, şifrelenmiş görüntü verisi iletim paketi içerisine eklenmek üzere Base64 kodlama şematiğine uygun olarak metin bazlı hale getirilmekte, eğer varsa görüntüdeki nesnelerin tespiti ile ilgili bilgiler ve şifre anahtarı, görüntü formatı ve dosya ismi, derin öğrenme modeliyle yorumlanan sahne ile ilgili sınıf etiketi gibi üst veri (metadata) halindeki bilgiler metin bazlı bir araya getirilerek JSON [29] paketi biçiminde sunucudan istekte bulunan istemcilere dağıtılmaktadır. Çalışmamızdaki deneylerimizde YOLO modeli kullanılarak sahnedeki nesnelerin sınıf etiketleri tespit edilmiş ve bu bir üst veri olarak JSON dosyası içerisinde kaydedilmiştir. Diğer bir doğrulama yöntemi olarak mevcut sahne için öğrenme sonucunda elde edilen temel yorumlamaya dayalı etiketleme yapmak için CLIP modeli kullanılarak yüksek doğrulukta nesne sınıfları tahminlenmiş, böylece bu bilgi bir üst veri olarak son kullanıcıya aktarılmıştır. JSON paketini alan istemci yine aynı anahtar ve bilgilere dayanarak görüntüyü deşifre etmekte ve deşifre görüntüdeki üst veri sonucunu son kullanıcıya doğrulama eşleşmesiyle göstermektedir.

Bu sistem literatürdeki diğer çalışmalardan farklı olarak mesaj iletiminde üst veri ile gönderilen doğrulama bilgisini kullanarak şifrelenmiş görüntünün doğru bir şekilde deşifre edildiğini garanti etmektedir. Sistem üç farklı veri kümesi üzerinden deneylerde gerçek zamanlı olarak iletişim modeli sayesinde denenmiştir. Kang ve Choi [15] çalışmasında kaotik bazlı ve yüksek güvenlikli bir sistem ile şifreleme yapılmasına rağmen bu çalışmamızda kullanılan Lehmer algoritmasının hesapsal karmaşıklığına göre Kang ve Choi çalışmasındaki karmaşıklık daha yüksek olmaktadır. Şifrelemeyi sadece ROI kullanarak yapan Kang ve Choi'nin çalışmasının aksine çalışmamızda tam ölçekli görüntülerle şifreleme ve deşifreleme yapılarak, sonucu doğrulama tümleşik olarak başarıyla gerçekleştirilmektedir. Yapılan çalışmada deneylerde elde edilen sonuçlar makalenin sonraki bölümlerinde detaylı olarak verilmektedir. Yapılan deneylerde Sunucu'dan İstemci'ye (Server to Client) her bir görüntünün şifrelenip mesaj paketine dönüştürülüp yollanması ve İstemci'nin aldığı mesaj paketindeki görüntüyü deşifreleyip doğrulaması ortalama olarak 50 Mbps hat üzerinden 24 milisaniye gecikme şartıyla toplamda uçtan uca 25 saniye sürmektedir.

5. Deneysel Sonuçlar (Experimental Results)

Bu bölümde deneylerimizde kullanılan veri kümelerinin detaylarına, başarım ölçütlerine ve değerlendirme sonuçlarına yer verilmektedir.

5.1. Veri kümeleri (Datasets)

Çalışmamızda kullanıcı deneyimi de göz önüne alınarak veri kümelerinde sınıf çeşitliliğinin artırılması ve böylece genel bir kullanım imkânı sağlanması hedeflenmiştir. Çalışmamızda üç adet veri kümesi kullanılmıştır. Bunlardan ilki, Massachussets Amherst Üniversitesi tarafından hazırlanmış ve halka açık kullanımla dağıtılan *Labeled Faces in the Wild* (LFW) veri kümesidir [30, 31]. İlgili veri kümesi, 5749 farklı insandan alınan görüntülerle oluşturulan toplamda 13233 insan yüzü görüntüsü içermektedir.

İkinci veri kümesi, *Oxford PASCAL VOC* 2012 [32-34] veri kümesidir. Halka açık kullanımla dağıtılmaktadır. Veri kümesinde arka plan (background) haricinde 20 farklı sınıf bulunmaktadır. Bu sınıflara ait etiketler şunlardır: insan, kedi, inek, koyun, köpek, at, kuş, uçak, bisiklet, tekne, otobüs, araba, motor, tren, şişe, yemek masası, saksı bitkisi, kanepe, sandalye, tv/monitör. Bu veri kümesinde, 11530 görüntü içerisinde 27450 adet işaretlenmiş nesne ve 6929 bölütleme verisi bulunmaktadır.

Üçüncü veri kümesi, *Unsegmented Sports News Videos* (CP Sports Minute, CPSM) [35, 36] veri kümesidir. Halka açık kullanımla dağıtılmaktadır. Veri kümesinde 10 farklı spor alanlarından oluşan sınıflar bulunmaktadır ve bu sınıflar şunlardır: yüzme, Amerikan futbolu, futbol, tenis, voleybol, beyzbol, golf, trekking, güreş. Veri kümesindeki videolar 60 ilâ 320 saniye uzunluğunda olabilen 240×320 çözünürlüğündeki 74 adet video parçasından oluşmakta ve bunlar üç ayrı spor branşını içermektedirler.

5.2. Başarım ölçütleri (Performance metrics)

Çalışmamızdaki veri kümelerinden alınan görüntüler ile yapılan veri şifreleme, deşifreleme ve iletimlerinde aşağıda bahsedilen ölçüm yöntemleri kullanılmış ve sonrasında sonuçları verilmiştir.

Yapısal Benzerlik Endeksi Ölçümü (Structural Similarity Index Measure, SSIM) [37], görüntülerin algılanan kalitesini sayısal olarak ölçebilmek için veya iki görüntü arasındaki benzerliği anlayabilmek için kullanılan bir birimdir. Bu ölçümde kontrast, parlaklık ve görüntü yapısı SSIM kontrast, parlaklık ve görüntü yapısını değerlendirerek bir sonuç çıkarır. SSIM, sıfır ile bir arasında değer alır. Değer bire yaklaştıkça görüntünün kalitesinin arttığı veya bir görüntü ile karşılaştırılırsa 1 değerinde görüntüyle aynı olduğu anlamına gelir. Tepe Sinyal Gürültü Oranı (Peak Signal Noise Ratio, PSNR) [38], görüntü kalitesini ölçmek için bilimsel araştırmalarda sık kullanılan bir ölçü birimidir. PSNR ölçütü göz önüne alındığında, örneğin, 8 bitlik görüntü derinliği olması halinde gürültü değerleri 30 ile 50 desibel (*dB*) arasında bir değer olmakta, 12 bitlik görüntülerde ise 60 *dB* veya daha yüksek değerlerde olabilmektedir. PSNR değeri yükseldikçe görüntünün kalitesinin arttığı ve parazitlerin azaldığı anlamına gelmektedir.

Denklem (3), Denklem (4) ve Denklem (5) sırasıyla Ortalama Kare Hatası (Mean Squared Error, MSE) değeri, PSNR ve SSIM değerlerinin hesaplanmasında kullanılan matematiksel formülleri vermektedir. Ortalama Kare Hatası tahmini ve gerçek değerler arasındaki farkı karesi alınmış hatanın ortalamasını alarak ölçer. Denklem (3)'te gürültüsüz bir $m \times n$ büyüklüğündeki I görüntüsü için K terimi gürültü yaklaşıklamasını ifade etmektedir. Denklem (4) ise MSE değeri tabanlı olarak PSNR değerini bulmak için hesaplanmaktadır. Buradaki MAX_I^2 terimi görüntüdeki olası maksimum piksel değerlerini ifade eder ve örnek başına 8 bit kullanıldığında bu değer 255 olur. Denklem (5)'te görüntüdeki genel büyüklüğü $N \times N$ olan denk boyuttaki iki pencere olarak verilen x ve y terimleri için benzerlik ölçüsüdür. Buradaki, μ_x terimi x görüntüsüne ait piksellerin, μ_y terimi y görüntüsüne ait piksel örneklem ortalamasıdır. σ_x^2 terimi x için varyansı, σ_y^2 terimi terimi y için varyansı ifade eder. Ayrıca, σ_{xy} terimi ise x ve y için kovaryansın hesaplamasıdır. SSIM formülündeki C_1 terimi incelendiğinde, $C_1=(k_1L)^2$ ve $C_2=(k_2L)^2$ zayıf payda ile bölünmeyi dengelemek için kullanılan iki değişkendir. Buna göre; L piksel değerlerinin dinamik aralığıdır. Buradaki k_1 terimi 0,01 değerini alırken, k_2 değeri ise 0,03 değerini almaktadır.

$$MSE = \frac{1}{mn}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}(I[i,j] - K[i,j])^2 \quad (3)$$

$$PSNR = 10log_{10}\left(\frac{MAX_I^2}{MSE}\right) \quad (4)$$

$$SSIM(x,y) = \frac{(2\mu_x\mu_y+C_1)(2\sigma_{xy}+C_2)}{(\mu_x^2+\mu_y^2+C_1)(\sigma_x^2+\sigma_y^2+C_2)} \quad (5)$$

Şifreleme ve deşifreleme kalitesine yönelik literatürdeki çalışmalarda kullanılan; Piksel Sayısı Değişim Hızı (Number of Pixels Change Rate, NPCR) [39, 40] ve Birleşik Ortalama Değişme Yoğunluğu (Unified Average Changing Intensity, UACI) [39, 40] şifreyi kırmaya yönelik diferansiyel ve lineer saldırıya dayanıklılık başarımını ve görüntü şifreleme yönteminin başarısını ölçmek için en sık kullanılan ölçütlerdir. İki görüntü arasındaki ilişkiyi belirleyen NPCR değeri göz önüne alındığında şifreleme öncesi ve şifreleme sonrası mevcut piksellerin değerleri birbirine eşitse 0, eşit değilse 1 değeri elde edilmektedir. Buna göre iki görüntü arasındaki fark ne kadar fazla ise NPCR değeri 1 değerine o kadar çok yaklaşmaktadır.

Düz metin (plain text) halindeki şifresiz orijinal görüntüdeki her bir pikselin uygulanan şifreleme işlemi sonrasındaki maruz kaldığı değişikliğe göre şifrelemenin kalitesini ölçebilmek için ölçütlerle değerlendirme yapılır. Buna göre şifresiz görüntüdeki bir piksellik değişikliği gösterecek şekilde şifreleme öncesi ve şifreleme sonrasında görüntülerin sırasıyla C^1 ve C^2 olduğu kabulüne göre, C^1 ve C^2 görüntülerindeki (i, j) piksel değerlerinin $C^1(i, j)$ ve $C^2(i, j)$ terimleri olarak alındığı formülde, $D(i, j)$ çift kutuplu dizisi üzerinden değişimin var olup olmadığına dair ölçümün tanımlandığını varsayalım. NPCR ve UACI ölçütlerinin değerleri bu bakış açısıyla hesaplanır. Buna dair matematiksel formüller aşağıdaki Denklem (6), Denklem (7) ve Denklem (8) ile verilmektedir. Bu formüllerdeki T terimi şifrelenmiş görüntüdeki toplam piksel sayısını, F terimi ise şifrelenmiş görüntü formatıyla uyumlu olan en büyük desteklenen piksel değerini ifade eder. Denklem (8)'de görüntüler arasındaki fark mutlak değer olarak elde edilmektedir. Denklem (7) ve Denklem (8) ise yüzdelik oran olarak verilmiştir. NPCR ölçütü, şifre kırmaya yönelik diferansiyel saldırılarda, şifreleme işlemi sırasında değeri değişmiş olan mutlak piksel sayısına odaklanmaktayken, UACI ölçütü ise birbirleriyle eşleştirilmiş sırasıyla C^1 ve C^2 olarak verilen iki görüntünün arasındaki ortalama farka odaklanmaktadır. NPCR ölçütünün alabildiği değerler [0,1] aralığında olmaktadır.

$$D(i,j) = \begin{cases} 0, & if\ C^1(i,j) = C^2(i,j) \\ 1, & if\ C^1(i,j) \neq C^2(i,j) \end{cases} \quad (6)$$

$$NPCR: \mathcal{N}(C^1, C^2) = \sum_{i,j} \frac{D(i,j)}{T} \times 100(\%) \quad (7)$$

$$UACI: \mathcal{U}(C^1, C^2) = \sum_{i,j} \frac{|C^1(i,j) - C^2(i,j)|}{F \cdot T} \times 100(\%) \quad (8)$$

Bilgi teorisinde Shannon tarafından tanımlanan entropi [41] ölçütü, bilgi belirsizliği ölçümü olarak kullanılır. Bilgi entropisi, belirli bir veriyi temsil etmek için gereken minimum ortalama bit'i ifade eder. Entropi, belirli bir görüntünün piksel değerlerinin olasılıksal dağılımından elde edilebilir. Çalışmamızdaki entropi değerleri verilen bir görüntüdeki piksellerin *X* dağılım fonksiyonuna uygun olmak üzere her bir pikselle ilişkili ortalama belirsizlik ölçütü Shannon entropisi olarak $H(X)$ terimiyle aşağıdaki Denklem (9)'a uygun olarak hesaplanmıştır. Elde edilen entropi değerleri görselleştirilerek ısı haritaları biçiminde bir sonraki bölümdeki şekillerde gösterilmektedir.

$$H(X) = -\sum_{i=0}^{2^{n-1}} p(x_i) log_2\, p(x_i) \quad (9)$$

Buradaki $p(x_i)$ terimi, *i* indisi olasılık kitle fonksiyon değeri için renk seviyelerinin sayısını göstermekte iken bu seviyelerle ilişkilendirilmiş olasılık değerini ifade etmektedir. Genel olarak 256×256 gri seviyesinde 8 bit formatında bir görüntüde (*n*=8), formüldeki $H(X)$ ile gösterilen entropinin maksimum değeri 8 olmaktadır. Bu maksimum entropi, piksellerden elde edilen önsel bilginin belirsiz olduğu, diğer bir deyişle bilginin "*yeterince karıştırılmış*" olduğu anlamına gelmektedir. Böylece, şifrelenen veri için entropi değeri 8 'e ne kadar yakın olursa, şifreleme işlemi o kadar iyi bir karıştırmaya sahip olmuş anlamına gelmektedir. Bu yolla, şifreleme sonrası orijinal veri hakkında doğrudan tahminleme yapılamayacak hale getirmek hedeflenmektedir. Bu nedenle, deşifreleme işleminde olduğu gibi görüntünün tekrar oluşturularak (reconstruction) orijinal görüntünün elde edilmesinde bu değerin minimize edilmesi ve böylece karmaşıklığın azaltılmasıyla orijinal bilgiye geri dönebilmek amaçlanmaktadır.

Bilgi elde etme kuramı (information theory) kapsamında sahne içerisindeki nesnelerin tespitine dair nesnel başarım ölçütleri [19, 32] incelendiğinde; bunlar arasında Doğruluk (Accuracy), Duyarlık (Precision), Anma (Recall), ölçütlerinin en temel ölçüm sonuçları olarak verildiği görülmektedir. Bu ölçütlerin detaylarına literatürdeki çalışmalarda yer verilmektedir [20, 32]. Ayrıca, Kesişim Birleşimi (Intersection over Union, *IoU*) ölçütü YOLO ve benzeri diğer modellerdeki tahminlenen nesnenin görüntüde kapladığı alanın tespitine dayanan bir ölçü değeri olarak, sınıf tahminlemesini yapan sinir ağı modeli tabanlı sınıflandırıcılarda kullanılan bir başarım ölçütüdür. Bu ölçüt, referans olarak alınan (ground-truth) gerçek sınırlayıcı kutu (bounding box) ile tahminleme yoluyla elde edilmiş sınırlayıcı kutunun kesişiminin görüntü üzerinde ne kadarlık bir alanı kapsadığını göstererek, bu kesişim alanı ile toplam birleşim alanın oranı olarak ifade etmektedir [32]. Genel Ortalama Duyarlık (mean Average Precision, *mAP*) ölçütü, nesne tespitindeki tahminlenen ilgili tüm sınıflar üzerinden alınan bu sınıflara ait Duyarlık değerleri ortalamalarının genel bir ortalaması şeklinde verilmektedir. Bu bahsi geçen ölçütlere temel olarak bilgi elde etmek kuramına ait bir sınıflandırma işlemindeki doğru pozitif (true positive, *DP*), doğru negatif (true negative, *DN*), yanlış pozitif (false positive, *YP*) ve yanlış negatif (false negative, *YN*) ölçüleri çalışmalarda nesne tespitinde sınıflandırma doğruluk oranının ve nesne tanıma oranı hesaplanması için kullanılmaktadır. Duyarlık ölçütü, tahmin edilen pozitif örneklerin kaçının gerçek pozitif örnekler olduğunu ifade etmektedir. Buna göre; sınıflandırma sonucunda DP değerinin tahminlemede elde edilen *DP* ile *YP* değerlerinin toplamına olan oranı olarak ifade edilir. Çalışmamızda YOLO modelinin deneylerdeki veri kümeleri için elde ettiği nesne tespiti tahminleme sonuçları *mAP* ölçütüyle ve *IoU* eşik değeri hem 0,5 hem de 0,5-0,95 aralığında alınarak hesaplanmış, böylece veri kümeleri üzerinden nesnel başarım değerlendirmesi yoluyla kıyaslanmıştır. *IuO* eşik değerinin buradaki anlamı, referans alınan sınırlayıcı kutu ile tahminleme yoluyla elde edilen sınırlayıcı kutunun kesişimi oranının 0,5 değerinden büyük olup olmadığına bakılarak eğer büyükse tahminlemenin doğru olarak kabul edilmesine dayanmaktadır.

5.2. Değerlendirme sonuçları (Evaluation results)

Çalışmamızda seçilen veri kümelerindeki görüntü verilerinin 256×256 çözünürlüğünde girdi ve çıktıları değerlendirilerek, şifrelemenin kalite yönünden başarısı, başarım ölçütlerine göre karşılaştırılarak Tablo 1'de ortalama değerler ve standart sapma değerleri olarak verilmiştir. Bu testler için Intel(R) Core(TM) i7-11370H model 3.30 GHz 64 bit işlemcili, 16 GB RAM bellekli, *n*Vidia *Geforce RTX* 3050 *Ti* 4 GB bellekli ekran kartlı ve Microsoft Windows 10 sürümü 64 bit işletim sistemine sahip bilgisayar üzerinde uygulamada Python v3.8 programlama dili kullanılmıştır. Derin öğrenme modelleri için Tensorflow v2.10.0 altyapısı [34] kullanılmıştır. Deneyler ölçüm sonuçlarını doğrulamak için üç kez tekrarlanmıştır. Tablo 1'de görülebileceği gibi çalışmamızdaki NPCR ölçütü ortalama değerlerinin tüm veri kümeleri için yaklaşık %99,61 olarak elde edilmiş olması ve tüm veri kümeleri için UACI ölçütü ortalamasının %14,70 değeriyle elde edilmiş olması, şifreleme işlemi sonrasında şifreli görüntü ile orijinal görüntü karşılaştırıldığında birbirlerine neredeyse hiç benzemedikleri anlamına gelmektedir. Tablo 1'de PSNR, SSIM, NPCR ve UACI değerlerinin üç veri kümesi üzerinden ortalama değerleri ve standart sapmaları görülmektedir.

Tablo 1. Deneylerdeki görüntü kalite ve rastgelelik tabanlı şifreleme başarım sonuçları (Image quality and randomness based encryption performance results in the experiments)

Veri kümesi	Betimleyici istatistik	PSNR	SSIM	NPCR(%)	UACI(%)
LFW	*Ortalama*	∞	1	99,60	19,50
	Standart sapma	0	0	0,0114	0,40
CPSM	*Ortalama*	∞	1	99,61	5,10
	Standart sapma	0	0	0,0457	0,05
PASCAL VOC 2012	*Ortalama*	∞	1	99,61	19,50
	Standart sapma	0	0	0,0069	0,57

Tablo 2'de Entropi değerlerinin üç veri kümesi üzerinden ortalama değerleri ve standart sapmaları görülmektedir. Tablo 2'deki entropi analizi ile elde edilmiş sonuçlara göre şifrelenmiş görüntüler bazında LFW ve PASCAL VOC 2012 veri kümelerinde dikkate alınan görüntü içerisindeki nesneler görüntünün büyük bir bölümünü homojen olarak kaplamaktadırlar. Bu nedenle ortalama entropi değerleri yaklaşık olarak birbirine yakın çıkmıştır. Aynı tabloda ortalama standart sapma değerleri ise oldukça küçük çıkarak görüntüde dikkate alınan nesnelerdeki renk dağılımlarına dair homojenitenin yüksek olduğunu göstermektedir. Tablo 2'deki şifrelenmiş görüntüler için en yüksek ortalama entropi değeri *CPSM* veri kümesi için 7,9999 olarak elde edilmiştir. Bu bakış açısıyla bu veri kümesi için çalışmamızda önerilen şifrelenme altyapısı sayesinde orijinal görüntüdeki piksel değerlerinin yeterince yüksek oranda karıştırılabildiği anlaşılmaktadır.

Tablo 2. Deneylerdeki şifrelenmiş ve deşifrelenmiş görüntüler için entropi analiz sonuçları (Entropy analysis results for encrypted and decrypted images in the experiments)

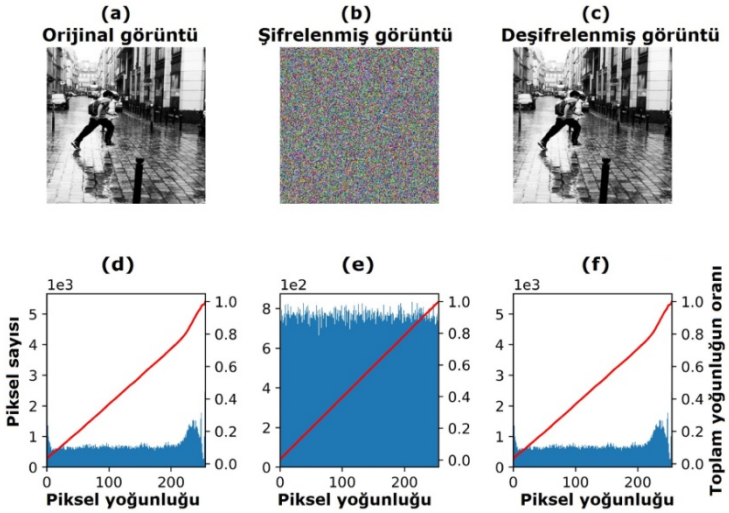
Veri kümesi	Betimleyici istatistik	Orijinal görüntü	Şifrelenmiş görüntü	Deşifrelenmiş görüntü
LFW	*Ortalama*	7,3988	7,9990	7,3988
	Standart sapma	0,3540	0,0001	0,3540
CPSM	*Ortalama*	7,5802	7,9999	7,5802
	Standart sapma	0,1509	0,0001	0,1509
PASCAL VOC 2012	*Ortalama*	7,4366	7,9991	7,4366
	Standart sapma	0,4447	0,0009	0,4447

MQTT bağlantısı ile uçtan uca iki bilgisayar arasında yapılan testlerde 50 Mbps hat üzerinden 24 milisaniye gecikmeyle, ilk veri gönderimi ağ trafiğine bağlı olarak ortalama 35 saniye ve diğer veriler ortalama 25 saniyede ulaştığı tespit edilmiştir. Gönderimlerde bahsedilen veri kümeleri kullanılarak yaklaşık 7000 görüntü verisi ile yapılan testte uçtan uca Gönderici ile Alıcı arasındaki mesaj iletim doğruluğu ve üretilen paketin tutarlılık kontrolü yapılmış, sonuçta tüm iletimdeki mesajlar üzerinden işlem %100 başarılı sonuçlanmıştır. Sonuç olarak gönderilen veri paketlerinin karşı tarafa hiçbir kayıp veya değişikliğe uğramadan ulaşmış olduğu garanti edilmiştir.

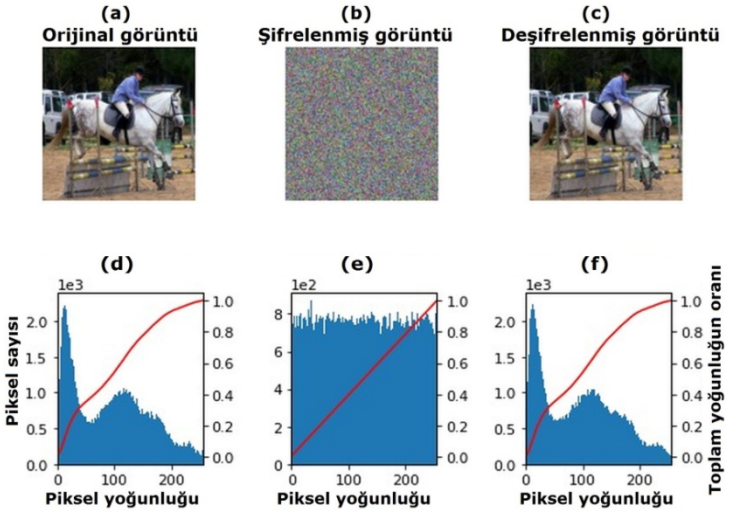
Çalışmamızda kullanılan YOLO modelinin CPSM, LFW ve PASCAL VOC 2012 veri kümelerindeki mevcut nesne ve sahne sınıflarına istinaden elde ettiği başarımlar olarak *IoU* eşik değeri 0,5 alınarak genel ortalama duyarlık (*mAP*) değerleri deneylerimizdeki veri kümeleri bazında sırasıyla; %51, %91,52 ve %73,12 olarak hesaplanmıştır. Ayrıca, *IoU* eşik değeri 0,5-0,95 aralığında alındığında LFW ve PASCAL VOC 2012 veri kümelerinde genel ortalama duyarlık değerleri sırasıyla; %68,21 ve %54,8'dir. CPSM veri kümesi ile yapılan nesne tespitinde kullanılan görüntüler taşıyıcı ana çerçeve (container frame) ile gösterildiğinden *IoU* eşik değerinin 0,5 olarak alınması başarım ölçümünde yeterli olmuştur. Literatürdeki Kang ve Choi [15] çalışması incelendiğinde şifrelemede kullanılan ROI bölgesinin ortalama olarak çözünürlüğü bizim çalışmamıza (256x256) en yakın olan (Kang ve Choi deneyindeki

"*Image3*" için) 173x592 piksel çözünürlüğünde toplam şifreleme zaman karmaşıklığı 2,2421 saniyedir. Çalışmamızda ise tam ölçekli görüntü kullanılarak yapılan şifrelemedeki toplam zaman karmaşıklığı 0,2638 saniyedir. Bu açıdan çalışmamızdaki zamansal karmaşıklığın oldukça düşük olması nedeniyle önerdiğimiz sistem son kullanıcı için daha tercih edilebilir bir sistemdir.

Aşağıdaki Şekil 3 ve Şekil 4'te *Oxford PASCAL VOC* 2012 veri kümesinden alınan görüntü örnekleri üzerinden şifreleme ve deşifreleme sonucunda görüntüdeki renk kanallarından elde edilen değerlerin ortalamalarının alınması ile hesaplanan görüntü olasılık yoğunluk fonksiyonuna göre histogram grafikleri verilirken, Şekil 5 ve Şekil 6'da ise şifrelenmemiş görüntülerin entropi değerlerine dair ısı haritası olarak verilmiştir. Şekil 3 ve Şekil 4'te (a) şıkkı orijinal görüntü, (b) şıkkı orijinal görüntünün şifrelenmiş hali, (c) şıkkı şifrelendikten sonra Alıcı tarafında şifresi çözülmüş görüntülerin örnekleri görülmektedir. Buradaki (a), (b) ve (c) şıkkındaki görüntülere ait sırasıyla ilgili histogram grafikleri bunların altındaki (d), (e) ve (f) şıklarında verilmiştir.



Şekil 3. Oxford PASCAL VOC veri kümesinden alınan birinci örnek görüntünün sonuçları (First sample image's results from the Oxford PASCAL VOC dataset)

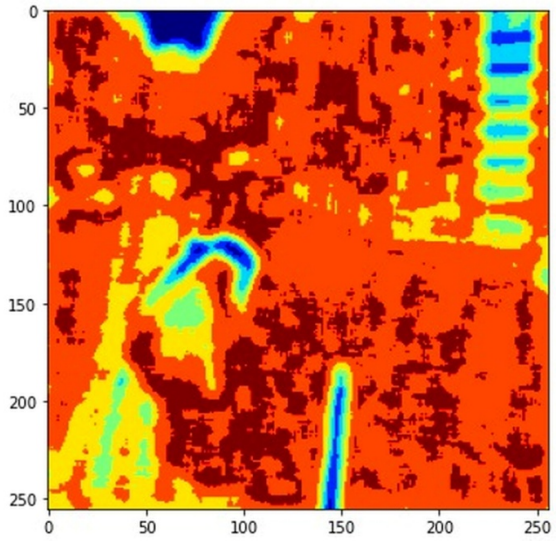


Şekil 4. Oxford PASCAL VOC 2012 veri kümesinden alınan ikinci örnek görüntünün sonuçları (Second sample image's results from the Oxford PASCAL VOC 2012 dataset)

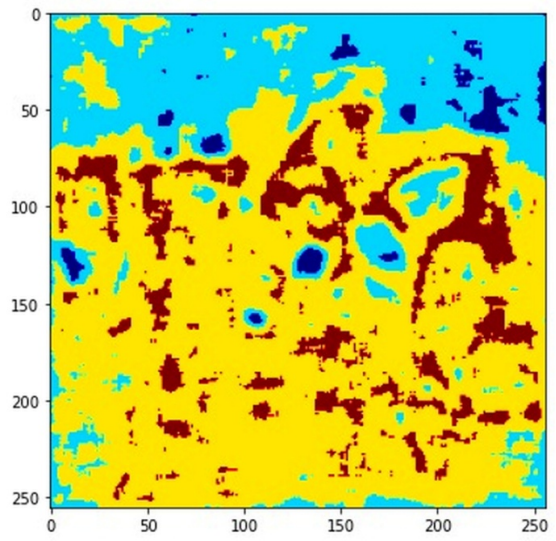
Şekil 3'den Şekil 12'ye kadar olan şekillerden görüleceği gibi şifreleme için kullanılan Gönderici taraftaki orijinal (şifresiz) görüntü ile Alıcı tarafta şifresi çözülmüş (deşifrelenmiş) görüntünün hem görsel olarak hem de olasılık yoğunluk fonksiyonuna göre histogram dağılımı olarak birebir aynı olduğu anlaşılmaktadır. Veri iletiminde bozulma olmaması, şifreleme ilâ deşifreleme altyapısının uçtan uca iletim sonucunda düzgün çalıştığı bu açıdan anlaşılmaktadır. Bu şekillerde şifrelenmiş görüntülerin histogramına dikkatlice bakıldığında tekdüze (uniform) bir dağılım olduğu, bu sayede tüm piksellerin renk aralığındaki tüm renklerle ifade edilerek karıştırılabilmiş olduğu ispatlanmaktadır. Buna göre,

ilgili Şekil 3, Şekil 4, Şekil 7, Şekil 9, Şekil 11 ve Şekil 14 şekillerindeki histogram dağılımlarının birikimli dağılım fonksiyonuna göre uygunluk derecesi kırmızı çizgi ile ortaya konulmaktadır. Aşağıda Şekil 5'te Şekil 3'teki, Şekil 6'da Şekil 4'teki, Şekil 8'de Şekil 7'deki ve Şekil 10 ise Şekil 9'daki görüntünün Alıcı tarafta deşifrelenmesi karşılığında oluşan sonuç görüntülerinden 10×10'luk çerçeve ile tanımlanmış çekirdekle (kernel) elde edilmiş entropi değerlerine dair ısı haritası verilmiştir. Buradaki iletim sonrası deşifrelenmiş görüntülerdeki entropi değerleri ve bunlara dair ısı haritaları iletim öncesi mevcut orijinal (şifresiz) görüntülerin entropi değerleri ile birebir aynı olmaktadır.

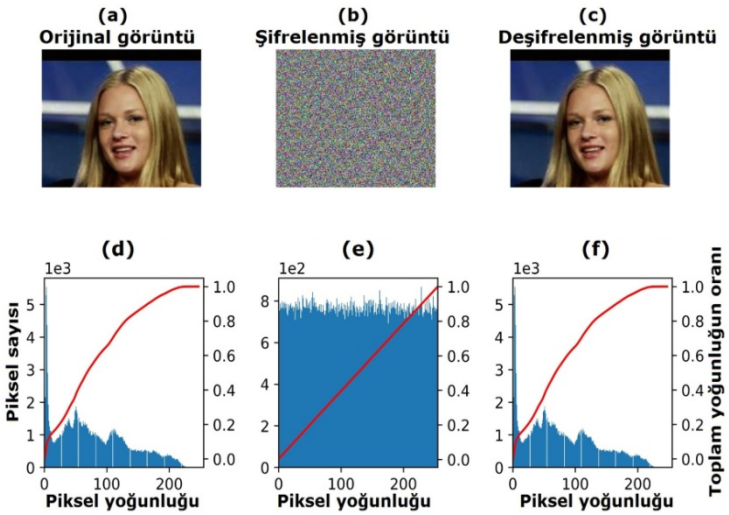
Görüntülerin şifreleme işlemi öncesi ve sonrasındaki durumlarını kıyaslamak adına görüntü entropisi [41], görüntünün belirli bir bölümündeki karmaşıklık seviyesini temsil eden bir değer elde etmede kullanılabilir. Bu karmaşıklık kaba kuvvet (brute force) saldırılarına karşı şifrelemenin ne kadar dayanıklı olduğuna dair bir ölçüm verir. Bir görüntünün entropi değeri kullanılarak buna dair üretilen histogram şifrelemenin saldırıya karşı dayanıklılığı (resilience) ve gürbüz (robust) olduğunu göstermektedir. Histogram, görüntüdeki farklı renk seviyesi olasılıklarını da göstermektedir.



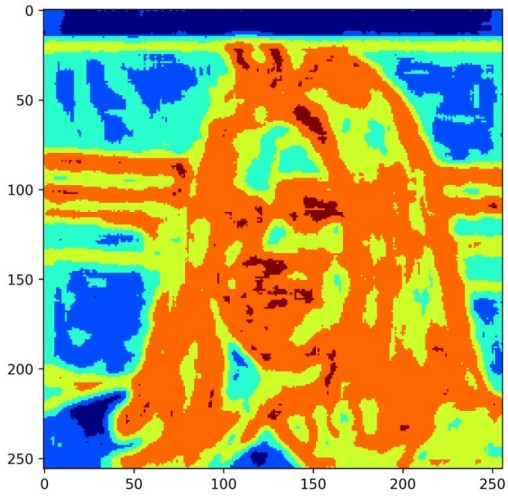
Şekil 5. Şekil 3'teki görüntünün Alıcı tarafta deşifrelenmesiyle alınan entropi sonuçlarının ısı haritası (Heat map for entropy results taken from the image in Figure 3 decrypted at Receiver side)



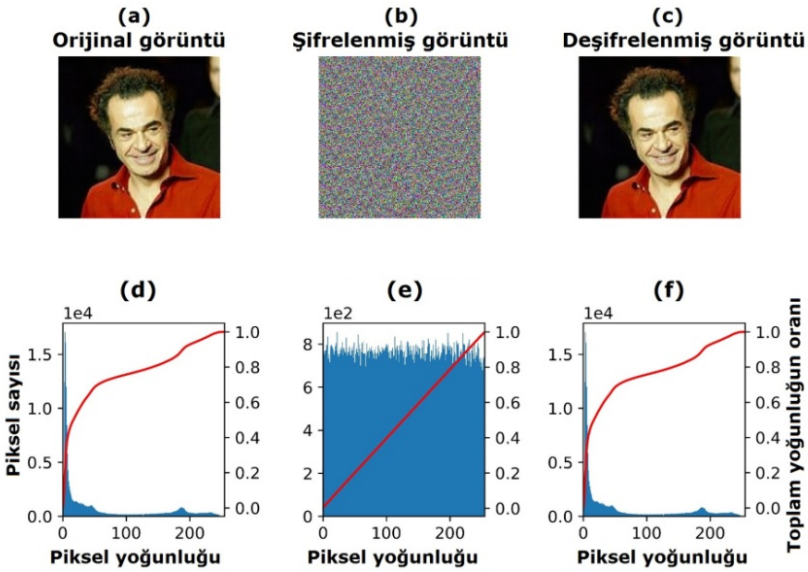
Şekil 6. Şekil 4'teki görüntünün Alıcı tarafta deşifrelenmesiyle alınan entropi sonuçlarının ısı haritası (Heat map for entropy results taken from the image in Figure 4 decrypted at Receiver side)



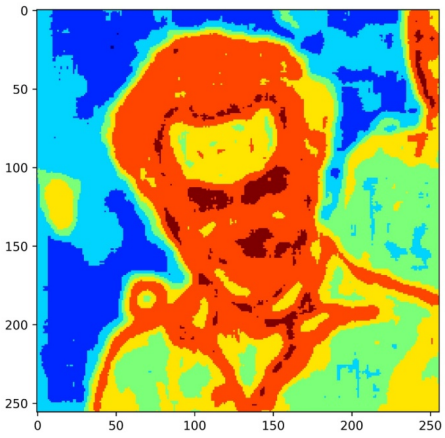
Şekil 7. Labeled Faces in the Wild (LFW) veri kümesinden alınan ilk örnek görüntünün sonuçları (First sample image's results from the Labeled Faces in the Wild dataset)



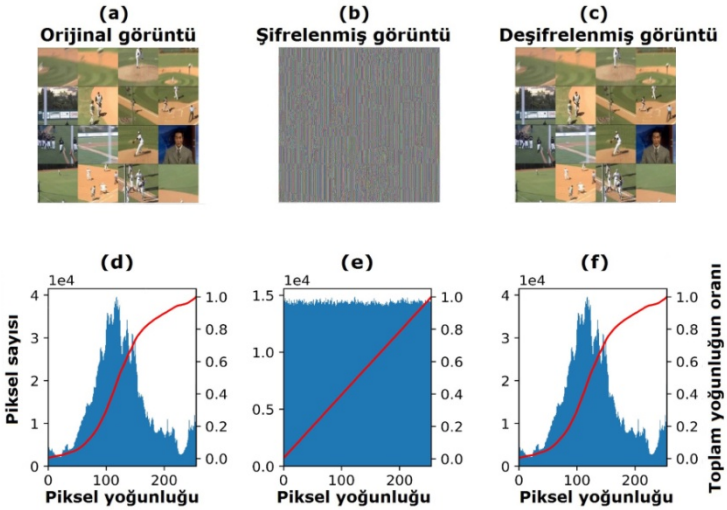
Şekil 8. Şekil 7'deki görüntünün Alıcı tarafta deşifrelenmesiyle alınan entropi sonuçlarının ısı haritası (Heat map for entropy results taken from the image in Figure 7 decrypted at Receiver side)



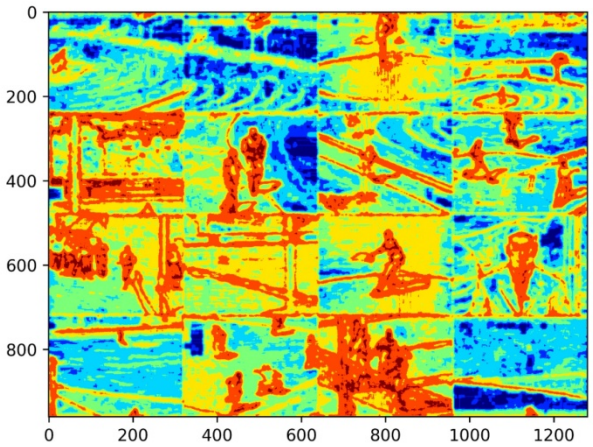
Şekil 9. Labeled Faces in the Wild (LFW) veri kümesinden alınan ikinci örnek görüntünün sonuçları (Second sample image's results from the Labeled Faces in the Wild dataset)



Şekil 10. Şekil 9'daki görüntünün Alıcı tarafta deşifrelenmesiyle alınan entropi sonuçlarının ısı haritası (Heat map for entropy results taken from the image in Figure 9 decrypted at Receiver side)



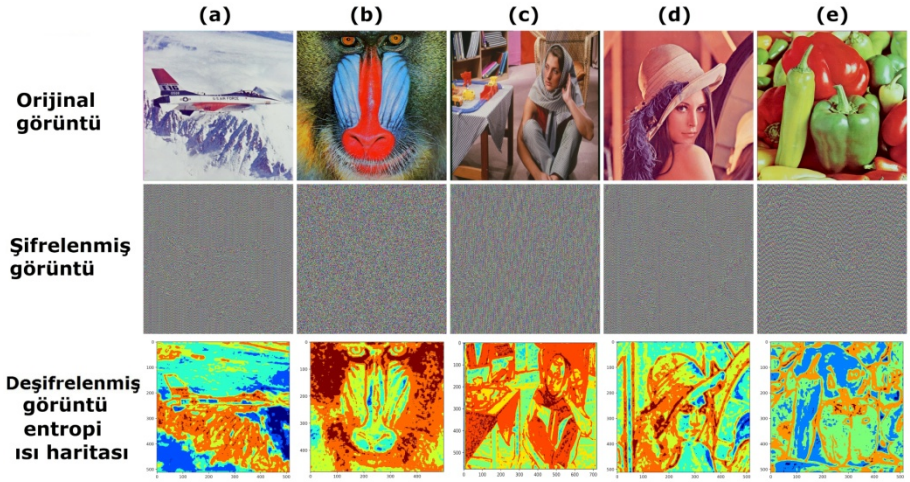
Şekil 11. Unsegmented Sports News Videos (CPSM) veri kümesinden alınan örnek taşıyıcı ana çerçeve görüntüsünün sonuçları (Sample container frame image's results from the Unsegmented Sports News Videos dataset)



Şekil 12. Şekil 11'deki görüntünün Alıcı tarafta deşifrelenmesiyle alınan entropi sonuçlarının ısı haritası (Heat map for entropy results taken from the image in Figure 11 decrypted at Receiver side)

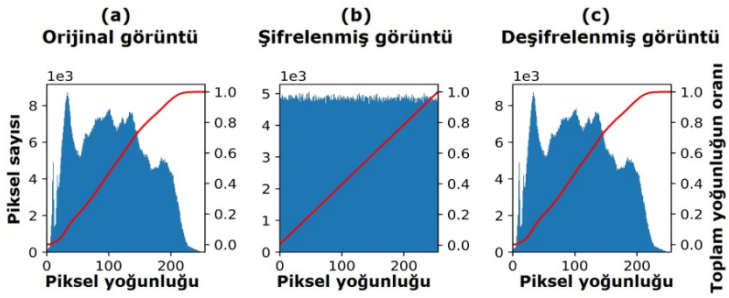
Bu bölümdeki entropi değerlerinin ısı haritalarından anlaşılacağı üzere, orijinal görüntülerde, birbirinden farklı bölgeler ile çeşitli nesneler ifade edilmektedir. Bu durum şifrelemenin karmaşıklığını arttırmaktadır. Bu nedenle entropi değerleri farklı veri kümeleri için farklı değerlerle elde edilmiştir. Şekil 13'te literatürdeki çalışmalarda sıklıkla kullanılan "*Air Plane*", "*Baboon*", "*Barbara*", "*Lenna*" ve "*Peppers*" test görüntüleri görülmektedir. Bazı çalışmalarda "*Lenna*" görüntüsü "*Lena*" şeklindeki

yazımı ile ifade edilmektedir. Şekil 13'te (a), (b), (c), (d) ve (e) şıkları için sırasıyla "*Air Plane*", "*Baboon*", "*Barbara*", "*Lenna*" ve "*Peppers*" görüntülerine ait olarak şekildeki satırlarda sırasıyla orijinal görüntü, şifrelenmiş görüntünün kendisi ve deşifrelenmiş görüntünün ise elde edilen entropi değerlerine dair ısı haritaları görülmektedir. Şekil 13'te verilmiş bu görüntüler için yüzdelik oran olarak NPCR değerleri sırasıyla; %99,6093 ile "*Air Plane*", %99,6093 ile "*Baboon*", %99,6095 ile "*Barbara*", %99,6093 ile "*Lenna*", %99,6093 ile "*Peppers*" olarak hesaplanmıştır. Yüzdelik oran olarak UACI değerleri ise sırasıyla; %9,86 ile "*Air Plane*", %10,39 ile "*Baboon*", %8,63 ile "*Barbara*", %9,77 ile "*Lenna*", %9,79 ile "*Peppers*" olarak hesaplanmıştır. Bu görüntüler için UACI ortalaması %9,68 olmaktadır. Deneylerdeki denemeler üzerinden alınan ortalama entropi değerlerine yakından bakıldığında, Şekil 13'te 256×256 çözünürlükte verilen orijinal test görüntüsünün çalışmamızdaki Lehmer tabanlı rastgele anahtar ile şifrelenmiş halinin entropi değerleri sırasıyla "*Air Plane*" görüntüsü için 7,9264, "*Baboon*" görüntüsü için 7,9341, "*Barbara*" görüntüsü için 7,9401, "*Lenna*" görüntüsü için 7,9303 ve "*Peppers*" görüntüsü için 7,9338 olarak elde edilmiştir. Mesaj iletimi sonrası deşifrelenmiş görüntünün ortalama entropi değerleri sırasıyla "*Air Plane*" görüntüsü için 6,7039, "*Baboon*" görüntüsü için 7,7853, "*Barbara*" görüntüsü için 7,7232, "*Lenna*" görüntüsü için 7,7520 ve "*Peppers*" görüntüsü için 7,7309 olarak elde edilmiştir.



Şekil 13. "*Air Plane*", "*Baboon*", "*Barbara*", "*Lenna*" ve "*Peppers*" görüntülerinin orijinal, şifrelenmiş ve Alıcı tarafta deşifrelenmiş görüntünün entropi değerleri ısı haritaları (Original, encrypted images and heat map for entropy results taken from the images decyrpyted at Receiver side as given as "*Air Plane*", "*Baboon*", "*Barbara*", "*Lenna*" and "*Peppers*")

Gönderici tarafa şifrelenmiş ve Alıcı tarafta deşifrelenmiş görüntülerdeki en yüksek entropi değerleri "*Barbara*" görüntüsü ile elde edilmiştir. Buna göre "*Barbara*" için şifrelenmiş haldeki görüntüde ortalama entropi değeri 7,9401 ve deşifrelenmiş haldeki görüntüde ise bu ölçüt 7,7232 olarak hesaplanmış, bu sayede en yüksek karıştırma oranına sahip şifrelemenin bu görüntü için oluştuğu gözlemlenmiştir. Tüm ortalama entropi değerleri Gönderici taraf ile Alıcı taraf arasında veri iletiminin ve şifreleme ilâ deşifreleme işleminin sorunsuz olarak gerçekleştirildiğini kanıtlamaktadır. Aşağıdaki Şekil 14'te deneylerimizdeki en yüksek entropi değerine sahip "*Barbara*" görüntüsüne ait olasılık yoğunluk fonksiyonuna göre histogram grafikleri şekildeki şıklar olarak sırasıyla; (a) orijinal görüntü, (b) şifrelenmiş görüntü ve (c) deşifrelenmiş görüntü için verilmiştir.



Şekil 14. "*Barbara*" görüntüsünün orijinal hali, şifrelenmiş hali ve deşifrelenmiş haline dair histogram grafikleri (Histogram plots for "*Barbara*" image as given as original image, encrypted image and decrypted image)

Şekil 14'ten görülebileceği gibi şifreleme ve deşifreleme sonucunda "*Barbara*" görüntüsünde hiçbir

bilgi kaybı yaşanmamış ve histogram ile ifade edilen görüntüye dair piksel yoğunluk dağılımları Alıcı taraftaki deşifreleme ile tekrar oluşturulan görüntüde başarılı bir biçimde aynı kalmıştır. Şifrelenen görüntünün histogramı ise orijinal görüntünün histogramına göre oldukça farklı elde edilmiştir. Şekildeki histogram dağılımlarının birikimli dağılım fonksiyonuna göre uygunluk derecesi kırmızı çizgi ile ortaya konulmaktadır. Çalışmamızda önerilen yöntemimiz sayesinde histogram tabanlı şifre kırma saldırılarına karşı yeterince dayanıklı olabilecek bir şifreleme işlemi yapıldığı anlaşılmaktadır. Aşağıdaki Tablo 3'de literatürdeki çeşitli çalışmalarla bizim çalışmamızda elde edilen deneysel sonuçlar karşılaştırılarak kıyaslanmaktadır. İlgili yayının atıfı altında deneylerinde kullandıkları görüntü çözünürlüğü de parantez içerisinde belirtilmiştir (200×200 ve 256×256 çözünürlükte olarak).

Tablo 3. Literatürdeki diğer çalışmalarla başarım kıyaslama tablosu (Performance benchmarking table for other's studies in the literature)

Görüntü	**Başarım ölçütleri**	**Pareek ve ark. [42] (*200x200*)**	**Zhu ve ark. [43] (*256x256*)**	**Çavuşoğlu ve ark. [44] (*256x256*)**	**Hanif ve ark. [45] (*256x256*)**	**Çalışmamız (*256x256*)**
	Entropi	-	-	-	7,9954	7,9264
Air Plane	*NPCR*	-	-	-	%99,65	%99,6093
	UACI	-	-	-	%33,8155	%9,86
	Entropi	7,9979	7,9968	7,9967	7,9952	7,9341
Baboon	*NPCR*	-	-	-	%99,6521	%99,6093
	UACI	-	-	-	%33,1627	%10,39
	Entropi	-	-	-	-	7,9401
Barbarra	*NPCR*	-	-	-	-	%99,6095
	UACI	-	-	-	-	%8,63
	Entropi	7,9996	7,9976	7,9958	7,9957	7,9303
Lenna	*NPCR*	-	%99,62	%0,0015	%99,5743	%99,6093
	UACI	-	%33,46	%0,0086	%33,0509	%9,77
	Entropi	7,9984	7,9975	7,9963	-	7,9338
Peppers	*NPCR*	-	-	-	-	%99,6093
	UACI	-	-	-	-	%9,79

Literatürdeki bu beş adet test görüntüsünü 200×200 ve 256×256 çözünürlükte olarak şifreleme deneylerinde kullanan çeşitli çalışmalara yakından bakıldığında; Pareek ve arkadaşları [42] çalışmasında, önerdikleri şifreleme yöntemiyle oluşturulan şifreli görüntünün entropi sonuçlarına göre sırasıyla "*Baboon*" görüntüsü için 7,9979, "*Lenna*" görüntüsü için 7,9996 ve "*Peppers*" görüntüsü için 7.9984 değerleri elde edilmiştir. Zhu ve arkadaşları [43] çalışmasında, önerilen kaos tabanlı *S* dönüşüm kutularına dayanan şifreleme yöntemiyle oluşturulan şifreli görüntünün entropi sonuçlarına göre sırasıyla "*Baboon*" görüntüsü için 7,9968, "*Lenna*" görüntüsü için 7,9976 ve "*Peppers*" görüntüsü için 7,9975 değerleri elde edilmiştir. Çavuşoğlu ve arkadaşları [44] çalışmasında, önerilen kaos tabanlı *S* dönüşüm kutuları kullanan şifreleme yöntemiyle oluşturulan şifreli görüntünün entropi sonuçlarına göre sırasıyla "*Baboon*" görüntüsü için 7,9967, "*Lenna*" görüntüsü için 7,9958 ve "*Peppers*" görüntüsü için 7,9963 değerleri elde edilmiştir. Hanif ve arkadaşları [45] çalışmasında, önerilen piksellerin blok seviyesi takası ve kaotik sistem tabanlı şifreleme yöntemiyle oluşturulan şifreli görüntünün entropi sonuçlarına göre sırasıyla "*Air Plane*" görüntüsü için 7,9954, "*Baboon*" görüntüsü için 7,9952, "*Lenna*" görüntüsü için 7,9957 değerleri elde edilmiştir. Hanif ve arkadaşlarının çalışmasında ayrıca NPRC ölçütü ile sırasıyla "*Air Plane*" görüntüsü için %99,6518, "*Baboon*" görüntüsü için %99,6521, "*Lenna*" görüntüsü için %99,5743 değerleri ve UACI ölçütü ile sırasıyla sırasıyla "*Air Plane*" görüntüsü için %33,8155, "*Baboon*" görüntüsü için %33,1627, "*Lenna*" görüntüsü için %33,0509 değerleri elde edilmiştir. Zhu ve arkadaşları çalışmasında "*Lenna*" görüntüsü için NPCR ölçütü ile %99,62 değeri ve UACI ölçütü ile %33,46 değeri elde edilmiştir. Çavuşoğlu ve arkadaşları çalışmasında ise hem NPCR hem de UACI değerleri sıfıra oldukça yakın çok küçük değerler olarak elde edilmiştir. Bu bahsi geçen çalışmalardaki tüm deneysel sonuçlara dair değerler incelendiğinde çalışmamızda önerilen Lehmer algoritmasıyla üretilen rastgele sayı tabanlı anahtarı kullanan şifreleme yöntemimizin literatürdeki çalışmalarla oldukça yakın sonuçlar ürettiği, görüntü ve videoların şifrelenerek güvenli bir biçimde iletiminde tercih edilebilir bir yaklaşım olduğu görülmektedir.

6. Sonuçlar ve Tartışma (Results and Discussion)

Bilgisayar ortamına aktarılan görüntü ve video, içerdikleri ham verinin iletilmesinde çeşitli algoritmalarla güvenliğin sağlanmasına uygun matris yapıları ve ilgili matematik altyapı sayesinde, zenginleştirilmiş bilgi taşıyan güvenli iletim ortamlarında sıklıkla kullanılmaktadır. Çalışmamızda, veri güvenliği ve görüntü veya video içerisindeki nesnelerin bilgisinin uçtan uca iletiminde şifreleme ve

deşifreleme mekanizması kurulurken, rastgele sayı üretimi tabanlı anahtar ile şifreleme için Lehmer algoritmasının kullanımı çeşitli veri kümeleri üzerinden denenmiştir. Bu denemelerde güvenli iletim ortamı kurulmasında MQTT protokolü sayesinde birbirini tamamlayan uç taraflar (end-to-end or peer-to-peer) için iletilen bilginin güvenilirliği, tekrar edilebilirliği, sürdürülebilirliği ve son kullanıcının bu bilgiyi elde ediş biçimi ön plana alınmıştır. Deneylerimizde başarımın ölçülmesinde şifrelemenin kalitesine dair başarım ölçütleri (NPCR ve UACI) yanı sıra deşifrelenmiş görüntünün orijinal görüntü ile arasındaki benzerliğe ve hata miktarına dayalı bozulma miktarı (SSIM ve PSNR) olup olmadığı da kontrol edilmiştir. Çalışmanın en önemli katkısı, Lehmer algoritmasının hem çeşitli ölçeklerde görüntüler için bit bazında kullanımının hem de derin öğrenme modelleri tabanlı (YOLO ve CLIP) nesne tespit ve sahne yorumlama yöntemleri ile harmanlanarak son kullanıcıya güvenilir bir bilişim sistemi altyapısı oluşturulmasında kullanılabilecek olmasıdır. İleriki çalışmalarımızda derin öğrenme modellerinin şifreleme mekanizmasına katkısı araştırılarak, Lehmer haricinde uygun rastgele sayı üreteci algoritmaların sisteme entegrasyonu ve uçtan uca şifrelemeye dair çeşitli veri kümeleri ile daha geniş kapsamlı bir çalışma yapılması planlanmaktadır.

Teşekkür (Acknowledgment)

Çalışmamızdaki deneylerde kullanılan halka açık erişimli görüntü veri kümeleri ve video veri kümesinin değerli sahiplerine teşekkür ederiz.

Çıkar Çatışması Beyanı (Conflict of Interest Statement)

Yazarlar tarafından herhangi bir çıkar çatışması bildirilmemiştir.

Kaynaklar (References)

[1] N. Koo, G. H. Cho, Byeonghwan and S. Kwon, "An Improvement of the Cipolla-Lehmer Type Algorithms", *National Institute for Mathematical Sciences*, arXiv Preprint: 1501.04036, 2015. doi:10.48550/arXiv.1501.04036

[2] D. Saif and A. Matrawy, "A Pure HTTP/3 Alternative to MQTT-over-QUIC in Resource-Constrained IoT", *Carleton University,* arXiv Preprint: 2106.12684, 2021. doi:10.48550/arXiv.2106.12684

[3] M. Ahmed and M. M. Akhtar, "Smart Home: Application using HTTP and MQTT as Communication Protocols*", Indian Institute of Technology,* arXiv Preprint: 2021.10339, Delhi, 2021. doi:10.48550/arXiv.2112.10339

[4] G. Perrone, M. Vecchio, J. D. Ser, F. Antonelli, V. Kapoor, "The Internet of Things: a Survey and Outlook*", Research Centre at eCampus University,* arXiv Preprint:1910.13965, Novedrate (Como), Italy, 2019. doi:10.48550/arXiv.1910.13965

[5] M. Lirzin and B. Markhoff, "Towards an ontology of HTTP interactions," 2020, France, arXiv Preprint: 2007.13475 [cs.AI], doi:10.48550/arXiv.2007.13475

[6] S. Hazelhurst, "A Proposal for Dynamic Access Lists for TCP/IP Packet Filering," 2001, *Programme for Highly Dependable Systems*, arXiv Preprint: cs/0110013. doi:10.48550/arXiv.cs/0110013

[7] S. Kumar, M. P. Andersen, H.-S. Kim and D. E. Culler, "Performant TCP for Low-Power Wireless Networks," 2018, *University of California, Berkeley,* arXiv Preprint: 1811.02721. doi:10.48550/arXiv.1811.02721

[8] Y. Bengio, A. C. Courville and P. Vincent, "Representation Learning: A Review and New Perspectives," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 8, pp. 1798-1828, Aug 2013. doi:10.1109/TPAMI.2013.50

[9] P. Jiang, D. Ergu, F. Liu, Y. Cai and B. Ma, "Review of Yolo Algorithm Developments," *Procedia Computer Science*, vol. 199, no. 1, pp. 1066-1073, 2022. doi:10.1016/j.procs.2022.01.135. ISSN 1877-0509

[10] A. Radford, J.-W. Kim, C. Hallacy, A. Ramesh, G. Goh, S. Agarwal, G. Sastry, A. Askell, P. Mishkin, J. Clark, G. Krueger and I. Sutskever, "Learning Transferable Visual Models From Natural Language Supervision," arXiv Preprint: 2103.00020, 2021. doi:10.48550/arXiv.2103.00020

[11] A.K. Banthia and N. Tiwari, "Image Encryption Using Pseudo Random Number Generators", *International Journal of Computer Applications*, vol. 67, no. 20, pp. 1-8, April 2013.

[12] V. Viswanatha, R. K. Chandana, A. C. Ramachandra, "Real Time Object Detection System with YOLO and CNN Models: A Review," arXiv Preprint: 2208.00773, 2022. doi:10.48550/arXiv.2208.00773

[13] H. Oğraş and M. R. Tür, "An Effective Image Encryption Algorithm Using Bit Reversal Permutation and a New Chaotic Map," *Gazi University Journal of Science*, vol. 35, no. 2, pp. 542-556, Jun. 2022. doi:10.35378/gujs.872818

[14] S. Somaraj and M.A. Hussain, "Securing Medical Images by Image Encryption using Key Image," *International Journal of Computer Applications*, vol. 104, no. 3, pp. 30-34, 2014. doi:10.5120/18184-9079

[15] S.-W. Kang, and U.-S., Choi, "ROI Image Encryption using YOLO and Chaotic Systems," *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 7, pp. 466-474, 2021. doi:10.14569/IJACSA.2021.0120754

[16] P. Lorek, G. Łoś, K. Gotfryd, and F. Zagórski, "On testing pseudorandom generators via statistical tests based on the arcsine law," arxiv Preprint: 1903.09805, 2019. doi:10.48550/arXiv.1903.09805

[17] K.S. Chua, "Chebyshev Polynomials And Higher Order Lucas Lehmer Algorithm," arXiv Preprint: 2010.02677, 2020. doi:10.48550/arXiv.2010.02677

[18] H.G. Katzgraber, "Random Numbers in Scientific Computing: An Introduction," Oldenburg, Germany, arXiv Preprint: 1005.4117, 2010. doi:10.48550/arXiv.1005.4117

[19] J. Redmon, S. Divvala, R. Girshick**,** and A. Farhadi, " You Only Look Once: Unified, Real-Time Object Detection," University of Washington, arXiv Preprint: 1506.02640, 2015. doi:10.48550/arXiv.1506.02640

[20] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A.C. Berg**,** and L. Fei-Fei, "ImageNet Large Scale Visual Recognition Challenge," Stanford University, Stanford, USA, arXiv Preprint: 1409.0575, 2014. doi:10.48550/arXiv.1409.0575

[21] T. Ridnik, H. Lawen, E. Ben-Baruch, and A. Noy, "Solving ImageNet: a Unified Scheme for Training any Backbone to Top Results," arXiv Preprint: 2204.03475, 2022. doi:10.48550/arXiv.2204.03475

[22] L. Beyer, O. J. Hénaff, A. Kolesnikov, X. Zhai and A. V. D. Oord, "Are we done with ImageNet?," arXiv Preprint: 2006.07159, 2020. doi:10.48550/arXiv.2006.07159

[23] Z. Wang, W. Liu, Q. He, X. Wu, and Z. Yi, "CLIP-GEN: Language-Free Training of a Text-to-Image Generator with CLIP," arXiv Preprint:2203.00386, 2022. doi:10.48550/arXiv.2203.00386

[24] Y. Cui, L. Zhao, F. Liang, Y. Li, and J. Shao, "Democratizing Contrastive Language-Image Pre-training: A CLIP Benchmark of Data, Model, and Supervision," arXiv Preprint: 2203.05796, 2022. doi:10.48550/arXiv.2203.05796

[25] H. You, L. Zhou, B. Xiao, N. Codella, Y. Cheng, R. Xu, S.-F. Chang, and L. Yuan, "Learning Visual Representation from Modality-Shared Contrastive Language-Image Pre-training," arXiv Preprint: 2207.12661, 2022. doi:10.48550/arXiv.2207.12661

[26] Katna, "Understanding katna," *katna.readthedocs.io*, 2022. [Online]. Available: https://katna.readthedocs.io/en/latest/understanding_katna.html. [Accessed: Aug., 2022].

[27] Python Imaging Library (PIL),"Pilow," *python-pillow.org*, 2022. [Online]. Available: https://python-pillow.org/. [Accessed: Aug., 2022].

[28] Y. Shi and Z. Zhang, "Communication Complexities of XOR functions," *Quantum Information & Computation*, vol. 9, no. 3, pp. 255–263, March 2009. doi:10.5555/2011781.2011786

[29] P. Bourhis, J. L. Reutter, F. Suárez and D. Vrgoč, "JSON: data model, query languages and schema specification", arXiv Preprint: 1701.02221, 2017. doi:10.48550/arXiv.1701.02221

[30] UMASS, "Labeled Faces in the Wild Home," viswww.cs.umass.edu, 2022. [Online]. Available: http://viswww.cs.umass.edu/lfw/. [Accessed: Aug., 2022].

[31] E. Learned-Miller, G.B. Huang, A. R, Chowdhury, H. Li, and G. Hua, "Labeled Faces in the Wild: A Survey," *In Advances in Face Detection and Facial Image Analysis*, pp. 189-248, 2016, doi:10.1007/978-3-319-25958-1_8

[32] M. Everingham, L.V. Gool, C.K.I. Williams, J. Winn and A. Zisserman, "The Pascal Visual Object Classes (VOC) Challenge," *International Journal of Computer Vision*, vol. 88, pp. 303–338, 2010. doi:10.1007/s11263-009-0275-4

[33] Oxford Pascal VOC, "Visual Object Classes Challenge 2012 (VOC2012)", *host.robots.ox.ac.uk*, 2022. [Online]. Available: http://host.robots.ox.ac.uk/pascal/VOC/voc2012/. [Accessed: Aug., 2022].

[34] Tensorflow, "PASCAL VOC dataset", *tensorflow.org,* 2022. [Online]. Available: https://www.tensorflow.org/datasets/catalog/voc. [Accessed: Aug., 2022].

[35] CPSM, "Unsegmented Sports News," *www.ed.ac.uk,* 2022. [Online]. Available: https://homepages.inf.ed.ac.uk/thospeda/downloads.html. [Accessed: Aug., 2022].

[36] T. M. Hospedales, S. Gong and T. Xiang, "Learning Tags from Unsegmented Videos of Multiple Human Actions," *2011 IEEE 11th International Conference on Data Mining*, pp. 251-259, 11-14 December 2011. doi: 10.1109/ICDM.2011.90

[37] J. Nilsson, and T. Akenine-Möller, "Understanding SSIM," arXiv Preprint: arXiv:2006.13846, 2020. doi: 10.48550/arXiv.2006.13846. [Accessed: Aug., 2022].

[38] O. Keleş, M.A. Yılmaz, A.M. Tekalp, C. Korkmaz, and Z. Doğan, "On the Computation of PSNR for a Set of Images or Video," arXiv Preprint: arXiv:2104.14868, 2021. doi:10.48550/arXiv.2104.14868

[39] A. A. Shah, A. Adeel, J. Ahmad, A. Al-Dubai, M. Gogate, A. Bishnu, M. Diyan, T. Hussain, K. Dashtipour, T. Ratnarajah, and A. Hussain, "A Novel Chaos-based Light-weight Image Encryption Scheme for Multi-modal Hearing Aids," arXiv Preprint: 2202.05662, 2022. doi: 10.48550/arXiv.2202.05662

[40] Y. Wu, J.P. Noonan and S. Agaian, "NPCR and UACI Randomness Tests for Image Encryption," *Cyber Journals: Multidisciplinary Journals in Science and Technology, Journal of Selected Areas in Telecommunications (JSAT)*, pp.31-38. 2011.

[41] D. J. C. MacKay, *Information Theory, Inference, and Learning Algorithms,* Cambridge University Press. UK, 2003.

[42] N. K. Pareek, V. Patidar, and K. K. Sud, "Colour Image Encryption Scheme Based on Permutation and Substitution Techniques," *In Advances in Computer Science and Information Technology, CCSIT 2011, Communications in Computer and Information Science*, vol. 131, pp. 413-327, 2011, doi:10.1007/978-3-642-17857-3_41

[43] S. Zhu, G. Wang, and C. Zhu, "A Secure and Fast Image Encryption Scheme based on Double Chaotic S-Boxes," *Entropy*, vol. 21, no. 8, p. 790, 2019. doi:10.3390/e21080790

[44] Ü. Çavuşoğlu, S. Kaçar, I. Pehlivan, and A. Zengin, "Secure image encryption algorithm design using a novel chaos based S-box," *Chaos Solitons Fractals*, vol. 95, pp. 92–101, 2017. doi:10.1016/j.chaos.2016.12.018

[45] M. Hanif, N. Iqbal, F.U. Rahman, M.A. Khan, T.M. Ghazal, S. Abbas, M. Ahmad, H.A. Hamadi, and C.Y. Yeun, "A Novel Grayscale Image Encryption Scheme Based on the Block-Level Swapping of Pixels and the Chaotic System," *Sensors*, vol. 22, no. 16, p. 6243, 2022. doi:10.3390/s22166243