ACTA INFOLOGICA 2022;6(1):127-139
dergipark.org.tr/acin

İSTANBUL UNIVERSITY PRESS

DOI: 10.26650/acin.1009226

RESEARCH ARTICLE

Görüntü İşleme Teknikleri ve Kelime Benzerliklerinin Ağırlıkları Algoritması Kullanılarak Geleneksel Sınavların Değerlendirilmesi

Grading Traditional Exams Using Image Processing Techniques and the Word Similarity Weights Algorithm

Yahya Gedik[1], **Serdar Solak**[2], **Mustafa Hikmet Bilgehan Uçar**[2]

[1] (Yüksek Lisans Öğrencisi), Kocaeli Üniversitesi, Bilişim Sistemleri Mühendisliği Bölümü, Kocaeli, Türkiye
[2] (Doç. Dr.), Kocaeli Üniversitesi, Bilişim Sistemleri Mühendisliği Bölümü, Kocaeli, Türkiye

ORCID: Y.G. 0000-0001-9014-251X;
S.S. 0000-0003-1081-1598;
M.H.B.U. 0000-0002-9023-0023

Corresponding author:
Serdar SOLAK
Kocaeli Üniversitesi, Bilişim Sistemleri Mühendisliği Bölümü, Kocaeli, Türkiye
E-mail address: serdars@kocaeli.edu.tr

Submitted: 14.10.2021
Revision Requested: 19.11.2021
Last Revision Received: 17.06.2022
Accepted: 19.06.2022
Published Online: 24.06.2022

Citation: Gedik, Y., Solak, S., & Ucar, M.B.H., (2022). Görüntü İşleme Teknikleri ve Kelime Benzerliklerinin Ağırlıkları Algoritması Kullanılarak Geleneksel Sınavların Değerlendirilmesi. *Acta Infologica, 6*(1), 127-139. https://doi.org/10.26650/acin.1009226

ÖZ

Teknolojide yaşanan hızlı gelişmeler eğitim alanına da yansımıştır. Eğitim kurumları, bilgisayar destekli olarak verdikleri eğitimleri çevrim içi veya çevrim dışı platformlarla desteklemeye başlamışlardır. Bu platformların kullanılması ile yapılan sınavların değerlendirilmesi sistem tarafından hızlı bir şekilde gerçekleştirilebilmektedir. Diğer taraftan, öğrencilere geleneksel yöntemler kullanılarak sınıf ortamında test, klasik, eşleştirmeli ve doğru yanlış tiplerinde soruları içeren sınavlar da uygulanmaktadır. Özellikle, sınavın sadece çoktan seçmeli sorulardan oluşmadığı durumlarda, değerlendirme işlemi uzun zaman almaktadır. Bu kapsamda yapılacak değerlendirmelerin doğru, güvenilir ve hızlı bir şekilde yürütülmesi tasarlanacak otomatikleşmiş akıllı uzman sistemlerle mümkündür. Bu çalışmada, sınıf ortamında gerçekleştirilen geleneksel sınavların hızlı ve güvenilir bir şekilde değerlendirilmesi için önerilen kelime benzerliklerinin ağırlıkları algoritması (KBAA) ve görüntü işleme tekniklerini kullanan bir yazılım geliştirilmiştir. Bu yazılım ile sınav kâğıtları üzerinde yer alan soru türlerinin ayrıştırılması, soru ve cevap ayrımı, el yazılarının tanımlanması ve cevapların değerlendirilmesi görüntü işleme ve KBAA kullanılarak gerçekleştirilmektedir. Deneysel çalışmalar sonucunda, sadece bir sınav kağıdının değerlendirilmesi saniyeler içerisinde gerçekleştirilirken sınavda yer alan tüm kağıtların toplu bir şekilde değerlendirilmesi ise değerlendirilen kâğıt başına yaklaşık 4 saniyede gerçekleştirilmektedir. Geliştirilen bu yazılım sayesinde geleneksel sınavların değerlendirilmesi hızlı, güvenilir, verimli ve adil bir şekilde gerçekleştirilebilmektedir.
Anahtar kelimeler: El Yazısı Tanıma, Görüntü İşleme, Optik Karakter Tanıma, Sınav Değerlendirme, Yapay Zekâ

ABSTRACT

The rapid developments in technology have also been reflected in education. Educational institutions have started to support their computer-aided trainings with online and offline platforms. Exam evaluations can be carried out quickly by the system using these platforms. Meanwhile, exams containing test, classical, matching, and true-false types of questions are also applied to students in the classroom environment using traditional methods. The evaluation process for these takes a long time, especially when an exam has more than just multiple-choice questions. In this context, carrying out evaluations to be made accurately, reliably, and quickly becomes possible with the intelligent expert systems being designed. Using the word similarity weights algorithm (WSWA) and image processing techniques, this study develops software for quickly and reliably evaluating the traditional exams held in classrooms. By means of the proposed software, question types are separated, answers are distinguished from questions, handwriting is identified, and answers are evaluated on the exam papers using image processing techniques and WSWA. As a result of the experimental studies, the evaluation of just one exam paper is carried out in seconds, while the collective evaluation of all the papers in the exam is carried out in approximately 4 seconds per paper being evaluated. Thanks to this software, traditional exams can be evaluated quickly, efficiently, and accurately.
Keywords: Image Processing, Artificial Intelligence, Handwriting Recognition, Optical Character Recognition, Exam Assessment

This work is licensed under Creative Commons Attribution-NonCommercial 4.0 International License

1. GİRİŞ

Eğitim kurumları, öğrencilerin ders başarımlarını değerlendirmek amacıyla çeşitli sınavlar yapmaktadır. Bu sınavlar, eğitmenlere eğitim süreçleri hakkında geri bildirim sağlamada kilit bir role sahiptir, bu nedenle hazırlanacak sınav sorularının türü ve seçiciliği oldukça önemlidir. Sınavlarda, farklı yöntemler ve ölçekler kullanılarak çeşitli tip, kategori ve kalitede sorular hazırlanmaktadır. Genellikle çoktan seçmeli sorulardan oluşan sınavlar hızlı, kolay, tutarlı, doğru ve daha objektif değerlendirilebilir olduğundan eğitim kurumları tarafından daha yaygın kullanılmaktadır (Yorke, 2009; Walvoord ve Anderson, 2011; Solak, Uçar ve Albadwieh, 2020). Özellikle bu sınavlar eğitmenler tarafından, sınav okuma zamanını ciddi bir şekilde azaltmasından dolayı tercih edilmektedir. Ayrıca son yıllarda kullanımı artan öğrenme yönetim sistemleri (Learning Management Systems, LMS) (McGill ve Klobas, 2009; Solak ve ark., 2016) sayesinde, eğitmenler uzaktan veya sınıf ortamında gerçekleştirilen çoktan seçmeli, eşleştirmeli, boşluk doldurma, doğru yanlış ve klasik sorulardan oluşan sınavları gerçek zamanlı olarak değerlendirip öğrencilerine ilan edebilmektedir. Ancak, LMS kullanmadan, geleneksel yöntemlerle sınıf ortamında yapılan ve içerisinde çoktan seçmeli, doğru yanlış, eşleştirme ve klasik soruları barındıran sınavların doğru ve etkili bir şekilde değerlendirilmesi, değerlendirici için yorucu bir süreç olup uzun zaman almaktadır.

Literatürde yapılan çalışmalar incelendiğinde, LMS sistemlerinin kullanılması ile yapılan çoktan seçmeli, doğru yanlış ve klasik sorulardan oluşan veya sınıf ortamında kâğıt ve kalem kullanılarak yapılan çoktan seçmeli sorulardan oluşan sınavların değerlendirilmesine yönelik olduğu görülmektedir. Sınıf ortamında kâğıt ve kalem kullanılarak yapılan çoktan seçmeli sınavların değerlendirilmesinde genellikle optik okuyucular kullanılmaktadır. Ancak, çoktan seçmeli sınavlarda doğrundan sınav kâğıdına işaretlenen cevapların değerlendirilebildiği bazı çalışmalarda mevcuttur (Tomić, Paunović ve Bosnić, 2020). Ayrıca sınıf ortamında yapılan kısa cevaplı soruların değerlendirilmesinde kelime benzerliği için yapay sinir ağları yöntemi de kullanılmaktadır (Chi ve Zhang, 2018; Clark ve ark., 2019).

Makalede, sınıf ortamında kâğıt ve kalem kullanılarak yapılan sınavların görüntü işleme teknikleri ve önerilen kelime benzerliklerinin ağırlıkları algoritması (KBAA) kullanılarak değerlendirilmesine yönelik çalışmalar gerçekleştirilmiştir. Sınıf ortamında gerçekleştirilecek bu sınavlarda, çoktan seçmeli, boşluk doldurma ve klasik sorular aynı anda sorulabilmektedir. Çalışmada ilk olarak, sınıf ortamında yapılan sınava ait kağıtlar kamera yardımıyla bilgisayar ortamına aktarılmaktadır. Daha sonra soruların tipleri ayrıştırılarak, öğrencilerin işaretlediği ve/veya yazmış olduğu cevaplar önerilen akıllı sistem tarafından tanınmaktadır. Sistem tarafından algılanan cevaplar daha önceden tanımlanan cevap anahtarı ile karşılaştırılarak değerlendirilmektedir. Sınıf ortamında yapılan sınavların değerlendirilmesinin zor olmasından dolayı, genellikle literatürde yapılan çoğu çalışma LMS sistemleri kullanılarak yapılan sınavların değerlendirilmesine odaklanmıştır. Önerilen çalışmanın katkısı, sınıf ortamında kâğıt ve kalem ile yapılan sınavların değerlendirilmesine için yeni bir çözüm yöntemi sunmaktadır. Ayrıca eğitimcilerin zaman kaybı yaşamadan, hızlı, etkili ve verimli bir şekilde bu sınavların değerlendirebilmesini sağlamaktır. Bununla birlikte, bu sınavlarda çoktan seçmeli, boşluk doldurma ve klasik soruların yer aldığı çok sayıda soru tipi aynı sınavda sorulabilmektedir. Çalışmanın bir başka katkısı ise, önerilen soru cevap ayrıştırma algoritması, anket çalışmalarının değerlendirilmesinde de kullanılabilir.

Makalenin ikinci bölümünde literatürde yapılan çalışmalar, üçüncü bölümde önerilen yöntem ve dördüncü bölümünde örnek bir çalışma sunulmaktadır. Son bölümde ise sonuçlar ve ileriki çalışma yer almaktadır.

2. LİTERATÜR TARAMASI

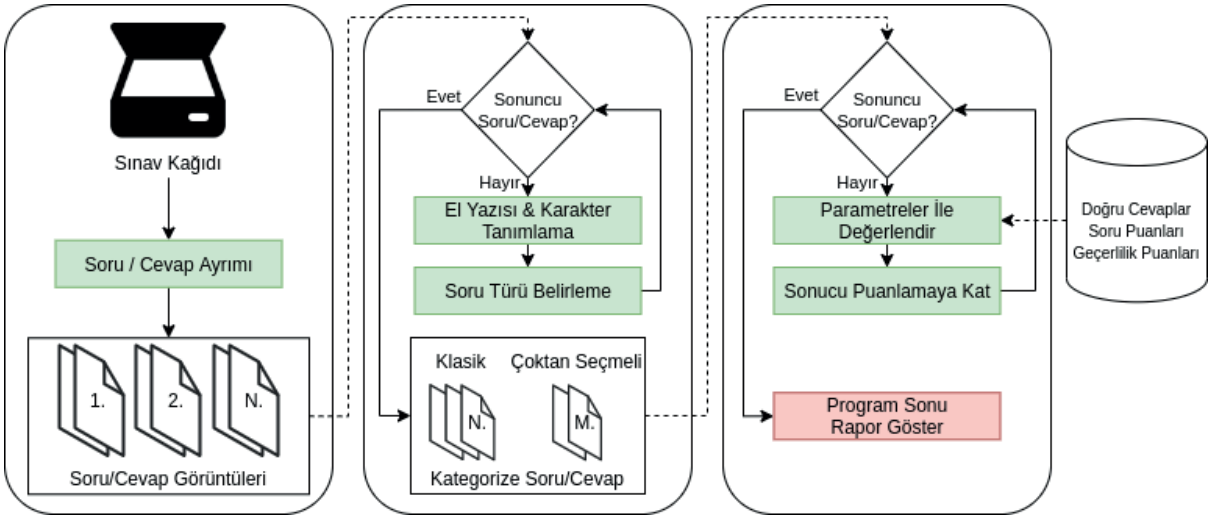
Literatürde yapılan çalışmalar incelendiğinde, sınıf ortamında yapılmış çoktan seçmeli sorulardan oluşan test sınavları, sınıf ortamında gerçekleştirilen klasik sınavlar ve LMS sistemleri kullanılarak gerçekleştirilen sınavlar için çeşitli değerlendirme sistemleri önerilmektedir. Bu kapsamda, 2010 yılında yapılan çalışmada, geleneksel kâğıt temelli sınavlar veya ev ödevlerinin dijital ortama aktarılarak, değerlendirici tarafından notlandırılmasını sağlayan bir sistem tasarlanmıştır (Bloomfield, 2010). 2012 yılında yapılan çalışmada ise, kâğıt tabanlı sınavların değerlendirilmesi amacıyla Moodle LMS'e eklenti olarak bir işaretleme aracı geliştirilmiştir (Villalon, 2012). 2013 yılında yapılan başka bir çalışmada ise, e-öğrenme değerlendirme sisteminin sağladığı avantajlar, kâğıt tabanlı sınavların değerlendirilmesi amacıyla kullanılmıştır. Bu çalışmada, geleneksel yöntemle sınava katılan öğrencilerin notlandırılması ve sınavlara ait analizleri dijital ortamda gerçekleşmektedir. Sınavlar,

geleneksel yöntemlerle yapıldıktan sonra, tarayıcı sayesinde internet üzerinden sunucuya aktarılmaktadır. Bu sunucuya herhangi bir bilgisayardan erişilerek de not verme işlemi yapılabilmektedir (Llamas-Nistal ve ark., 2013). 2016 yılında yapılan çalışmada, Eclipse editörüne bir eklenti hazırlanarak programlama eğitimi için bir değerlendirme sistemi geliştirilmiştir. Bu sistem Java programlama dilinin açıklama özelliğini kullanarak öğrencilerin yazmış olduğu hatalı kod parçalarının tespit edilmesini sağlamakta ve aynı zamanda bu programlama dilinin yansıma uygulama programlama ara-yüzü (Application Programming Interface, API) kullanarak doğruluk tespitini yapmaktadır (Nemeth ve Tejfel, 2016). 2016 yılında yapılan bir başka çalışmada ise elektronik ortamda yapılan bilgisayar programlama derslerinin değerlendirmesi için sistem önerilmiştir. Öğrencilerin değerlendirilmesinde, eğitimcilere çeşitli analizler yapma olanağı sağlayarak geri bildirimlerin sağlanması gerçekleştirilmiştir (Rajala ve ark., 2016). 2019 yılında yapılan çalışmada, çevrimiçi olarak gerçekleştirilen klasik sınavların değerlendirilmesi için bir sistem geliştirilmiştir. Çalışma, ikili yaklaşımlar yöntemini kullanarak, öğrencilerin yazmış olduğu cevapları anlamsal olarak değerlendirmekte ve puanlayabilmektedir (Tashu, Esclamado ve Horvath, 2019). 2019 yılında yapılan çalışmada, kısa cevaplı soruları puanlamak için otomatik değerlendirme sistemi önerilmektedir. Çalışmada öğrenci yanıtları ile doğru yanıtlar arasında anlamsal ilişki kurularak model çıkarılmaktadır. Önermiş oldukları bu modeli K-12 veri seti üzerinde deneyerek iyi bir performans elde ettiklerini öne sürmüşlerdir (Liu ve ark., 2019). 2020 yılında çoktan seçmeli sınavların optik karakter formu kullanılmadan değerlendirilmesine yönelik bir çalışma gerçekleştirilmiştir. Bu çalışmada sınıf ortamında kâğıt ve kalem kullanarak gerçekleştirilen çoktan seçmeli sınavlar için hibrit bir sınav oluşturma ve değerlendirme sistemi önerilmektedir (Tomić, Paunović ve Bosnić, 2020). 2020 yılında yapılan bir çalışmada kısa cevaplı sorulara verilen yanıtlar otomatik değerlendiren bir sistem önermekte ve sınava girenlerin yeterliliğini tahmin etme sorununu ele almak için NLP ve psikometrik ölçüm alanlarındaki yaklaşımları bir araya getirmektedir. Çalışmada sunulan yaklaşım, incelenen kişinin yeterliliğini doğrudan modeller ve eğitim verisini oluşturmak için etiketlenmiş verilere ihtiyaç duymamaktadır (Ha ve ark., 2020). 2020 yılında yapılan çalışmada, öğrencilerin cevap kâğıtlarına yazdığı cevap ile doğru cevap arasındaki metin benzerliğini bulmak için WordNet grafı kullanılmaktadır. Yazarlar çalışmasında, cevap metninin anlamsal ilişkilerini dikkate alabilen yeni bir işaretleme algoritması sunulmaktadır (Vij, Tayal ve Jain, 2020). 2021 yılında yapılan çalışmada eğik karakterleri tanımada karşılaşılan zorluklara dikkat çekilmiş ve yeni bir yöntem önerilmiştir. Önerilen bu yöntemde açık kaynak kodlu Tesseract kütüphanesinin eğik karakter tanıma başarısı Titreşim (Dithering) görüntü işleme algoritması kullanılarak arttırılmıştır (Çelik 2021). 2021 yılında yapılan çalışmada, makine öğrenmesi algoritmaları kullanılarak Portekizce yazılan açık uçlu soruları değerlendirmek amacıyla bir sistem önerilmektedir. Sistem sayesinde alınan verileri ön işleme tabi tuttuktan sonra Lojistik Regresyon, Ekstra Ağaçlar, Rastgele Orman ve Çok Katmanlı Algılayıcı algoritmalarının hem çok sınıflı hem de ikili sınıflandırma için iyi sonuçlar verdiğini gözlemlemişlerdir (Gomez Rocha ve ark., 2021). 2021 yılında yapılan çalışmada kağıt kalem kullanılarak yapılan klasik sınavların değerlendirilmesi için otomatik değerlendirme sistemi önerilmektedir. Öğrencilerin yazdığı cevapların tanınması için optik karakter okuma, değerlendirilmesi için makine öğrenmesi / doğal dil işleme kullanılmaktadır (Sanuvala ve Fatima, 2021). 2021 yılında yapılan bir çalışmada, bilgisayar programlama derslerinde, programlama kodunun analiz edilerek açıklamaların yazılacağı kısa cevaplı soruların otomatik değerlendirilmesi için doğal dil işleme tabanlı bir sistem önerilmiştir (Fowler ve ark., 2021).

3. ÖNERİLEN YÖNTEM

Sınıf ortamında kâğıt ve kalem kullanılarak yapılan geleneksel klasik sınavların bilgisayar destekli olarak değerlendirilmesindeki en büyük sorunlardan birisi sınav kağıdının bilgisayar tarafından anlaşılır hale getirilmesidir. Bu sorunu çözerken karşılaşılan temel problemlerden biri kişiden kişiye değişen el yazılarının tanınmasıdır. El yazılarının tanınması (Handwriting Recognition, HWR) için bulut tabanlı ve yerel olmak üzere iki çözüm yöntemi bulunmaktadır (Plamondon ve Srihari, 2000). Bulut tabanlı HWR büyük şirketlerin bulut hizmetleri tarafından sağlanan, yüksek doğruluk oranına sahip ve hızlı olan bir el yazısı tanıma servisidir (Gao ve ark., 2011; Shaukat ve ark., 2020). Bu servislerin maliyeti düşük olmasına rağmen kullanım oranına bağlı olarak maliyetler artmaktadır. Ayrıca bulut tabanlı HWR servisinin kullanılması için kesintisiz internet bağlantısı gerekmektedir. Yerel HWR ise el yazısını tespit edecek bir yapay zekânın eğitilmesi ile mümkün olmaktadır (Marti ve Bunke, 2002; Plötz ve Fink, 2009). Bu yöntem bulut tabanlı el yazısı tespitine göre daha düşük oranda doğruluk vermektedir ancak maliyeti düşük ve çevrimdışı çalışma özelliği bulunmaktadır. Ayrıca, OpenCV (Bradski ve Kaehler, 2008) ve Tesseract (Smith, 2007) gibi çeşitli açık kaynak kodlu görüntü işleme kütüphaneleri de görüntülerde yer alan karakterlerin ve el yazılarının dijitalleştirilmesi sağlamaktadır (Çelik, 2020; Şahin ve ark., 2022).

Çalışmada, kâğıt ve kalem kullanılarak sınıf ortamında yapılan sınavları değerlendirmek amacıyla görüntü işleme ve yapay zekâ teknikleri kullanılarak yeni bir yöntem önerilmektedir. Şekil 1'de önerilen çalışmada kullanılan yöntemin akış şeması sunulmaktadır. Görüldüğü üzere, ilk olarak sisteme bağlı olan bir kamera veya tarayıcı vasıtasıyla sınav kağıdındaki veriler bilgisayar ortamına aktarılmaktadır. Daha sonra, görüntüsü alınan sınav kâğıdının karakteristik bir desen yardımıyla soru ve cevap yerlerinin tespit edilmesi ve ayrıştırılması yapılmaktadır. Bu işlemden sonra, bulut tabanlı HWR kullanılarak el yazısı tanıma işlemi gerçekleştirilmektedir. El yazısı tanıma işleminden sonra, öğrencinin verdiği cevaplar, cevap anahtarı ile karşılaştırılarak sonuçlar değerlendirilmektedir. Sonuç elde edilirken ilk olarak soru tipi bulunmaktadır. Soru tipinin çoktan seçmeli olması durumunda, işaretli cevap olup olmadığı ve cevabın doğruluğu kontrol edilmektedir. Soru tipinin klasik olması durumunda cevabın doğruluğu, öğrencinin vermiş olduğu cevap ile doğru cevap arasındaki kelime ağırlıklarına göre benzeşmeyi hesaplayan bir algoritma yardımı değerlendirilmektedir.



Şekil 1. Önerilen yönteme ait çalışma akış şeması.

Çalışmada C++ ve Python programlama dilleri ve OpenCV kütüphanesi kullanılmıştır. Geliştirilen sistemin tüm işletim sistemlerinde çalışması için web tabanlı ara-yüz hazırlanmış olup C# programlama dili kullanılarak yazılmıştır. Makale kapsamında gerçekleştirilen yazılım, sınav dokümanının hazırlanması ve bilgisayara ortamına aktarımı, karakter ve el yazısı tanıma, cevapların değerlendirilmesi, sonuçların elde edilmesi ve web ara-yüzünün geliştirilmesi olmak üzere beş alt bölümde ele alınmaktadır.

3.1. Sınav Belgesinin Hazırlanması ve Bilgisayara Aktarımı

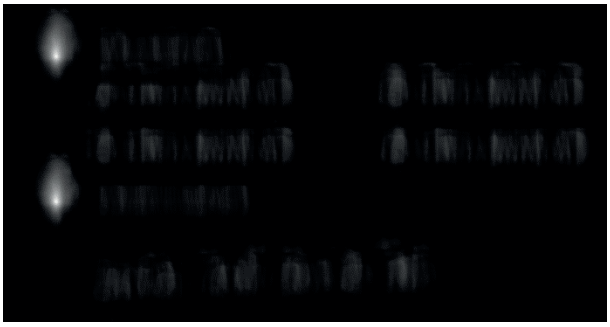
Sınıf ortamında kalem ve kâğıt kullanılarak yapılan geleneksel sınavların değerlendirilebilmesi için sınav dokümanının belirli bir formata uygun olarak hazırlanması gerekmektedir. Bu kapsamda gerçekleştirilecek sınavlar için hazırlanacak belgeye ait format, sınav kağıdına olabildiğince az ekleme yapılarak tasarlanmıştır. Oluşturulacak sınav kâğıdında yer alan her sorunun başına bir desen yerleştirilmesi gerekmektedir. Yerleştirilen bu desen sorunun ve cevabın kâğıt üzerindeki konumlarını hesaplanmak için kullanılmaktadır.



(a) (b)

Şekil 2 Soruların başına yerleştirilmiş olan (a) desen örneği ve (b) maske görüntüsü

Sınav kâğıdı üzerindeki soru ve cevapların konumları belirlemek için kullanılan desen ve ona ait maske görüntüleri sırasıyla Şekil 2.(a) ve Şekil 2.(b)'de yer almaktadır. Çalışma kapsamında seçilen ve Şekil 2.(a)'da yer alan desen, örnek olarak seçilmiş olup farklı bir desen görüntüsü de kullanılabilmektedir. Seçilen desenin sınav kâğıdı üzerindeki herhangi bir görüntüye kıyasla daha fazla özelliğe sahip olması gerekmektedir. Şekil 2.(a)'da örnek olarak sunulan desen, renk tonları, kenar ve köşe gibi oldukça fazla özelliğe sahip olmasından dolayı soruları tespit etme konusunda hata oranını minimize etmektedir. Desen eşleştirme işlemini sadece renkli desen görüntüsü ile eşleştirmek mümkündür, ancak bu görüntüde belirgin özelliklerin görüldüğü kısımlar dışında etrafında beyaz alanlar bulunmaktadır. Bu alanlara bazen harflerin gelmesi desen eşleştirmeyi yanıltmaktadır. Bu doğrultuda, Şekil 2.(b)'de sunulduğu gibi bir maske görüntüsü hazırlanması gerekmektedir. Bu maske görüntüsü desen görüntüsündeki belirgin alanları beyaz renkle, alakasız kısımları ise siyah renkle göstermektedir. Kâğıt, desen ve maske görüntüleri kullanılarak desen eşleştirme işlemi gerçekleştirilir. Desen eşleştirme işlemi sonucunda Şekil 3'te sunulan monokrom bir görüntü elde edilmektedir.



Şekil 3. Desen eşleştirmeden sonra elde edilen sonuç görüntüsü

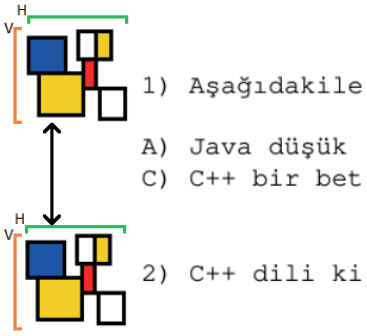
Şekil 3'te sunulan, görüntünün solunda yer alan beyaz pikseller eşleşmenin daha belirgin olduğu kısımlardır. Bu kısımlar şu anki haliyle bulanıktır. Bu noktaların daha belirgin hale getirilebilmesi için öncelikle eşikleme işlemi uygulanması gerekmektedir. Eşikleme işlemi sonrasında ikili görüntü elde edilmektedir. Eşikleme işleminde kullanılan eşik değeri fonksiyonu, koşul sağlanması durumunda beyaz rengi, sağlamadığı durumda ise siyah rengi vermektedir. Şekil 4 eşikleme işleminden sonra elde edilen ikili görüntüyü temsil etmektedir. Şekil 4'te yer alan görüntü kullanılarak, soru ve cevaplarının yerlerinin hesaplanması mümkündür. Soru ve cevapların konumlarının bulunabilmesi için sırasıyla matematiksel minimize ve maksimize işlemi ve boşlukların doldurulması işlemi bir döngü içinde gerçekleştirilmektedir.



Şekil 4. Eşikleme işlemi uygulanmasından sonra oluşan ikili görüntü

Matematiksel minimize ve maksimize işlemi, görüntüdeki en düşük ve en büyük değerli piksellerin konumlarının bulunmasını sağlamaktadır. Eşikleme işlemi sonunda elde edilen ikili görüntü sadece siyah ve beyaz piksellerden oluştuğu için elde edilecek maksimum değer 255 ve minimum değer 0 olacaktır. Bu işlem sonunda en büyük değerli pikselin konumu elde edileceğinden dolayı bulunan ilk konum, eşleşmenin ilk gerçekleştiği konum olacaktır. Bu konum ayrıca ilk sorunun ve cevabın sayfadaki yüksekliğini belirlemektedir. Döngü tekrar etmeden önce maksimum değerin bulunduğu konum için boşluk doldurma işlemi gerçekleştirilmektedir. Bu işlem belirlenen pikseli ve etrafındaki aynı renkteki pikselleri belirlenen renge boyamak için kullanılır. Boşluk doldurma işlemi ile siyaha dönüştürülen beyaz pikseller bir sonraki döngüde matematiksel

minimize ve maksimize fonksiyonu tarafından fark edilmeyecektir. Bu sayede ilk soru için yapılan tüm işlemler bir sonraki soru ve cevaplar için gerçekleştirilebilir. Her döngüde soruların ve cevapların kâğıt üzerindeki konumlarının belirlenmesi için basit cebirsel işlemler uygulanmaktadır.



Şekil 5. Soru ve cevapların yerleri hesaplanırken desenler arasındaki pozisyon ilişkisi

Şekil 5'te iki soru ve onlara ait cevap alanlarına dair bazı bilgiler sunulmaktadır. Matematiksel minimize ve maksimize işleminden sonra elde edilen en yüksek değerli pikselin konumu, görüntüdeki desenin başlangıcını belirtmektedir. Bu bilgiye dayanarak sorunun yüksekliğini hesaplamak için desen görüntüsünün yüksekliği ve elde edilen en yüksek değerli pikselin yüksekliği toplanır (V+Max). Sorunun genişliği tespit edilirken, eşleşmenin gerçekleştiği konumdan kâğıdın en sağına kadar olan kısım alınmaktadır. Kağıttaki cevap alanlarının bulunması için daha önce bir soru alanının hesaplanmış olması gereklidir, yani cevapların konumlarının bulunması 2. sorudan itibaren başlar. 2. soru için sorunun yüksekliği ve genişliği hesaplandıktan sonra bir önceki cevap alanının yüksekliği ve genişliğinin hesaplanması gerçekleştirilir. Cevap alanının başlangıcını, önceki sorunun bitiş yüksekliği belirlemektedir. Cevap alanının bitiş yüksekliği için o anki sorunun başlangıç yüksekliği kullanılmaktadır. Cevap alanının genişliği içinse kâğıdın en solundan en sağına kadar olan kısım alınmaktadır. En sonda yer alan cevap alanı için önceki sorunun bitiş yüksekliğinden kâğıdın bitişine kadar olan kısım alınır, genişliği ise diğer cevap alanları ile aynı şekilde hesaplanmaktadır.

Sorunun yer aldığı alanın yüksekliğinin hesaplanması için aşağıdaki eşitlik kullanılır:

$$H_q = [H_m, (H_m + H_t)] \tag{1}$$

Eşitlik (1)'de, H_m eşleşmenin gerçekleştiği yüksekliği, H_t desen yüksekliğini ve H_q elde edilecek olan soru alanının yüksekliğini belirtmektedir. Eşleşme yüksekliği ve desen yüksekliği toplanarak sorunun yer aldığı alanın yüksekliği hesaplanmıştır. Cevabın yer aldığı alanın yüksekliğinin hesaplanması için aşağıdaki eşitlik kullanılır;

$$H_a = \left[\left(H_p + H_t\right), H_m\right] \tag{2}$$

Eşitlik (2)'de, H_p önceki sorunun kapladığı alanın yüksekliğini, H_t desen yüksekliğini, H_m o anki soru eşleşme yüksekliğini ve H_a ise elde edilecek olan cevap alanın yüksekliğini belirtmektedir. Önceki sorunun yer aldığı alanın bitiş yüksekliği ile mevcut sorunun yer aldığı alanın başlangıç yüksekliği arasında kalan alan, cevap alanını karşılık gelmektedir.

3.2. Karakter ve El Yazısı Tanıma

Sınav belgesinin hazırlanması ve bilgisayara aktarılması işleminden sonrasında sınav kağıtlarından elde edilen meta veriler karakter ve el yazısı tanıma aşamasında kullanılmaktadır. Bu aşamada ilk olarak soruların türleri belirlenmektedir. Eğer klasik veya boşluk doldurma soruları varsa bulut tabanlı HWR kullanılarak el yazısı tanıma işlemi gerçekleştirilmektedir. Çalışma kapsamında, Google tarafından sağlanan yapay zekâ destekli, yüksek doğruluğa ve hıza sahip Google bulut görme (Google Cloud Vision) HWR teknolojisi (Liu ve ark., 2020) kullanılmaktadır.

Şekil 6'da bozuk kenarlar ve kötü el yazısı gibi hatalara sahip olan örnek görüntü bilgisayar ortamında oluşturulmuştur. Çalışmada kullanılan bulut tabanlı HWR servisinin doğruluğunu tespit etmek amacıyla test aşamasında kullanılmıştır. Şekil 6'dan da görüldüğü üzere sistem yazıyı doğru bir şekilde algılamıştır.



Şekil 6. Fare kullanılarak yazılmış bir el yazısının (çerçeve içinde) tespiti (mor).

Önerilen çalışma, klasik soruların yanı sıra çoktan seçmeli sorular için yüksek doğrulukta tespit gerçekleştirerek değerlendirme yapabilmektedir. Çoktan seçmeli soruların tespiti için aşağıda verilen ön koşulların sağlanması gerekmektedir.

- Klasik sorular için yerleştirilen desen çoktan seçmeli soruların başında da bulunmalıdır.
- Çoktan seçmeli sorular 'A)', 'B)', 'C)', 'D)' seçeneklerini içerebilir, seçenek sayısı artırılabilir ancak seçeneklerin formatı değişmemelidir.
- Cevap seçenekleri arasında en az iki veya üç karakterlik boşluk bulunmalıdır.
- İşaretleme yapıldığında seçenek harfi okunamayacak kadar işaretli olmalıdır.

Belirtilen ön koşulların sağlanmaması durumunda çoktan seçmeli soruların değerlendirilmesinde problem yaşanabilmektedir. Çalışmanın çoktan seçmeli sorulardan oluşan sınavlarda kullanılan optik formlara göre avantajı, ek bir kâğıt ihtiyacının olmamasıdır. Ayrıca seçeneklerin cevap alanı için ayrılan alanda herhangi bir sırada ve yerde olması önemli bir avantajdır. Sınav kağıdında yer alan klasik ve çoktan seçmeli sorular bu işlemler sonunda sade yazı formatında bir dosyaya içerisine kaydedilmektedir. Bu dosya içerisinde, sorunun türünü ve verilen cevabın ne olduğunu bilgisi saklanmaktadır. Bu işlem, sınava katılan tüm öğrencilere ait sınav kağıtlarında yer alan tüm sorular için uygulanmaktadır.

3.3. Sorulara Verilen Yanıtların Değerlendirilmesi

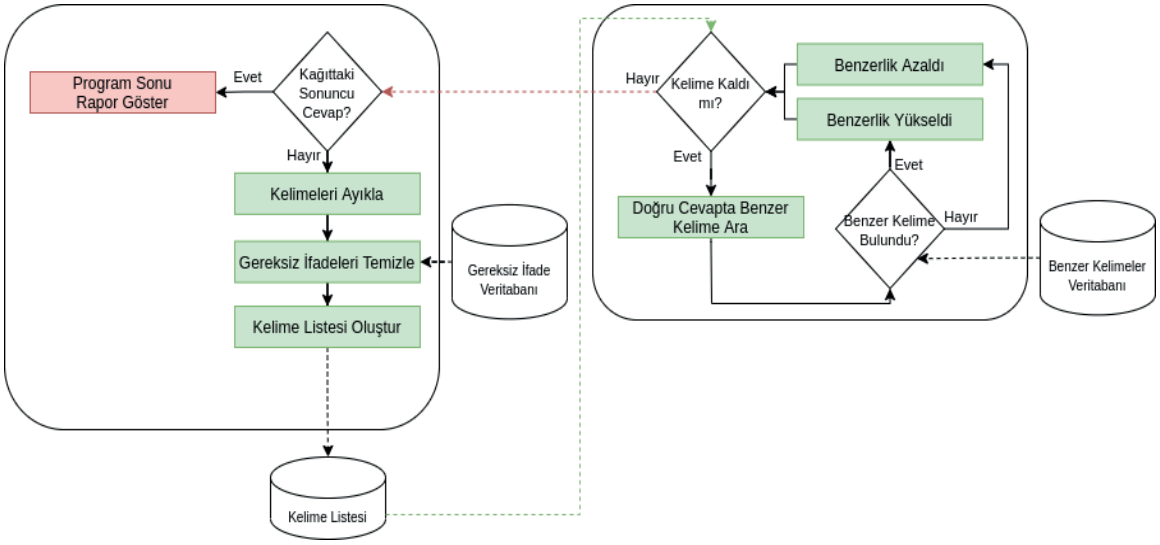
Karakter ve el yazısı tanıma aşaması sonunda sınav kağıtları için elde edilen dosyalar cevap değerlendirme aşamasında kullanılmaktadır. Klasik ve çoktan seçmeli sorulara verilen yanıtların değerlendirilmesinde literatürde GPT-2 (Generative Pre-trained Transformer) (Lagler ve ark., 2013; Lee ve Hsiang, 2020) modelini kullanan yapay zeka teknikleri bulunmaktadır. GPT-2, OpenAI şirketi tarafından hazırlanmış bir açık kaynak kodlu yapay zekâ modelidir. Bu model önceden eğitilmiş olup belirli bir amaca hizmet etmek için kullanılabilmektedir. Örneğin, GPT-2 George Orwell'in 1984 kitabından bir cümle okuyup devamını üretmeye çalışabilir veya bir yöneticinin ağzından bir cümle oluşturabilmektedir. GPT-2'nin Türkçe değerlendirme yapabilmesi için, eğitim aşamasında birçok Türkçe kelime ve cümleye hâkim olması gerekmektedir. Klasik sınav sorularının değerlendirilebilmesi için, GPT-2 modelinin Şekil 7'de sunulduğu gibi bir veri seti ile eğitilmesi gerekmektedir.

```
A: Dünyanın yarıçapı kaç kilometredir?
B: 6371.
A: Pi sayısı kaçtır?
B: 3.14159265359.
A: Ford şirketi kaç yılında kurulmuştur?
```

Şekil 7. GPT-2 için sağlanabilecek veri setinin bir kısmı.

Bu veri seti GPT-2 modelinin soru-cevap ve cevap-cevap ilişkilerini öğrenmesini amaçlamaktadır. Fakat böyle bir veri setinin faydalı olması ve iyi sonuçlar vermesi için oldukça büyük olması gerekmektedir. Ayrıca, eğitim için oldukça fazla zaman harcanması gerekir. GPT-2 yapay zekâ modelinin sonuç oluşturması oldukça uzun sürmektedir. Gerçekleştirilen çalışmadaki amaçlardan biri her sınav kâğıdı için sonuç alma süresini minimize etmektir. Bu çalışmada, klasik soruların değerlendirme işleminin daha hızlı gerçekleştirilebilmesi için kelime benzerliklerinin ağırlıkları isimli özel bir değerlendirme algoritması tasarlanmıştır. Bu algoritma, iki cümle arasındaki benzerliği kelime ağırlıklarına göre benzeşmeyi hesaplayarak bulmaktadır. Bu hesaplama işlemi sınav kâğıdı üzerindeki cevap ile doğru cevabı karşılaştırmakta ve soruya verilen yanıtın doğru olup olmadığını bulmaktadır. Bu işlemde yaygın olarak karakter ve dizi işlemleri kullanıldığından, hızlı ve verimli bir performans elde etmek amacıyla Python programlama dili kullanılmaktadır. Şekil 8'de klasik soruların değerlendirilmesi

için önerilen benzerlik hesaplama algoritması sunulmaktadır. Algoritmada, cevapların değerlendirilmesine geçilmeden önce el yazısı tanıma bölümünde elde edilen veriler üzerinde temizleme işlemleri yapılmaktadır.



Şekil 8. Klasik soruların benzerliğinin hesaplanmasında kullanılan algoritma.

Veri temizleme sürecinde aşağıdaki işlemler sırasıyla gerçekleştirilmektedir;

- Noktalama işaretlerini ayrıştırılıp silinmesi.
- Sayı ve yazıları ayrılması.
- Bazı doldurucu (örneğin, ‘ve’, ‘veya’, ‘gibi’) ifadeleri atlanması
- Eklerin ihmal edilmesi (örneğin, ‘-de’, ‘-da’, ‘-lar’, ‘ler’)

Veri temizleme sürecinden sonra kelime sayım işlemi gerçekleştirilmektedir. Değerlendirme için öncelikle benzerlik verisini içeren bir dosya okunmaktadır. Bu dosya içerisinde birbirine benzeyen ifadeler gruplandırılmış bir şekilde yer almaktadır. Bu dosyada bulunamayan ifadeler kendi başına sayılır. Daha sonra kelimeler birbirine ne kadar benzediklerine göre bir ‘benzerlik’ değeri oluşturulmaktadır. Bu değer doğru cevap ve kâğıt üzerindeki cevabın ne kadar birbirine benzediğini belirtmektedir. Eğitimci veya sınav değerlendirmesinden sorumlu kişi değerlendirme yapılmadan önce sınav tanımlama ara-yüzü yardımıyla her soru için kontrol değeri ve puan belirlemesi gerçekleştirebilmektedir. Bu ara-yüz sayesinde soru türleri de belirtilebilmektedir. Soru tipleri belirtilmese de değerlendirme yapılabilmektedir ancak daha sağlıklı ve doğru sonuçlar elde etmek için belirtilmesi tavsiye edilmektedir. Kontrol değeri ile değerlendirilen cevaplar raporlanmak için bir Json (Javascript Object Notation) dosyasına kaydedilmektedir.

3.4. Değerlendirme Sonrası Sonuçların Elde Edilmesi

Sorulara verilen yanıtların değerlendirilmesinden sonra içerisinde sonuçların yer aldığı Json dosyası elde edilmektedir. Bu dosya içerisinde WEB-tabanlı ara-yüzünde kullanılacak olan önceki aşamalarda hesaplanmış anahtar veriler bulunmaktadır. Bu Json dosyasının içerisinde, type, result, weight, answer ve diff alanları mevcuttur. Bu alanlara ait tanımlamalar;

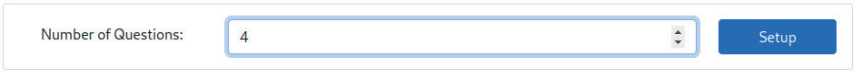
- ‘type’ alanı, sorunun türüne dair bilgi içermektedir. Bu alan, sorunun klasik ya da çoktan seçmeli olmasına göre sırasıyla ‘multi’ ya da ‘classic’ değerlerini alır.
- ‘result’ alanı, verilen cevabın doğru cevaba ne kadar yakın olduğunu belirtmektedir. Eğer soru çoktan seçmeli ise sonuç 1 veya 0 değerini almaktadır.
- ‘weight’ alanı, sorunun puanını belirlemektedir. Bu alana ait değer, eğitimci veya değerlendirmeyi yapacak sorumlu kişi tarafından sınav değerlendirme aşamasından önce hazırlanan sınav tanımlama ara-yüzünde belirtilmektedir.

- ‘answer’ alanı, soruya ait sınav kâğıdı üzerinde verilen cevabın tamamını içermektedir. Bu cevap üzerinde henüz temizleme gerçekleşmemiştir.
- ‘diff’ alanı, cevabın doğru kabul edilebilmesi için gerekli olan kontrol değeridir. Bu değer çoktan seçmeli sorular için önemli değildir ve ‘-1’ olarak belirlenmektedir. Klasik sorular için hesaplanan benzerlik değeri, bu değerin üzerindeyse cevap doğru kabul edilmektedir.

Json dosyası, web ara-yüzü tarafından otomatik olarak okunmaktadır. Ayrıca sınav kağıtları için genel puanlamaları gösteren kullanıcı dostu bir ara-yüz ile sonuçlar sunulmaktadır. Bu ara-yüzde her soru için doğruluk oranları ve puanlamalar gösterilmektedir.

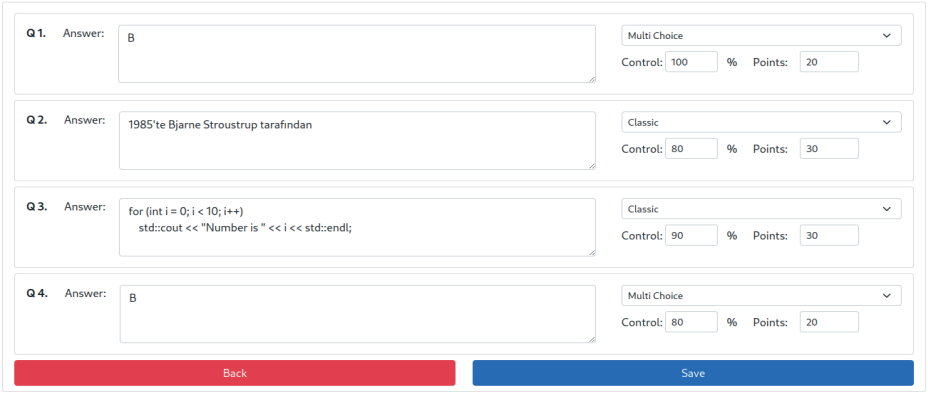
3.5. WEB-Tabanlı Sınav Değerlendirme Ara-yüzü

Önerilen sistemin kullanılmasını kolaylaştırmak amacıyla WEB-tabanlı sınav hazırlık ve değerlendirme uygulama ara-yüzü geliştirilmiştir. Geliştirilen bu WEB-tabanlı uygulama ara-yüzü için .NET platformunu kullanan özel bir sunucu geliştirilmiştir. Bu sunucu, WEB ve masaüstü ortamlarda çalışabilmektedir. Uygulama başlatıldığı zaman kullanıcılar herhangi bir WEB tarayıcısı üzerinden uygulamanın sunulduğu adrese ulaşarak sistemi kullanabilmektedirler. Şekil 9’da sunulduğu üzere, öncelikli olarak değerlendirilecek sınavında yer alacak soru sayısı belirlenir.



Şekil 9. Soru sayısı belirleme ekranı.

Şekil 10’da sunulan ikinci aşamada ise, kullanıcının sınav için cevap anahtarı oluşturması gerekmektedir. Her sorunun başında numaralandırma yer almaktadır. Bu numara sınavdaki soru numarası ile eşleşmesi gereklidir. Her soru için ayrıca bir cevap alanı bulunmaktadır. Bu alana klasik soru için doğru cevabı yazılırken, çoktan seçmeli sorular için tek harf girilmesi yeterlidir.

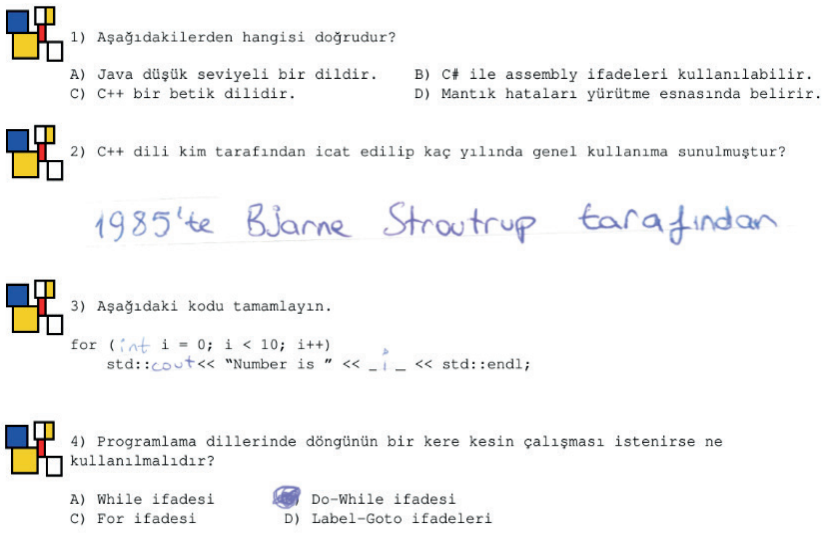


Şekil 10. Cevap anahtarı hazırlama ekranı.

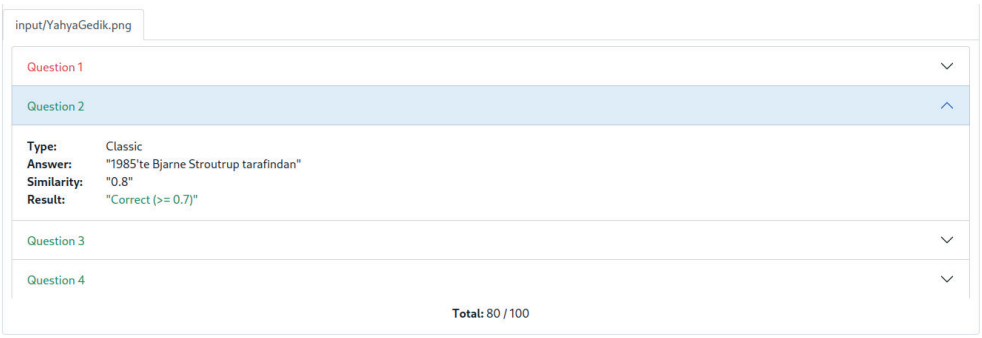
WEB ara-yüzünde, cevap alanının karşısında soru türü seçim alanı bulunmaktadır. Normal şartlarda uygulama, soru türlerini otomatik olarak tespit etmektedir. Ancak soru türünün bu bölümde belirtilmesi ileride gerçekleşebilecek tanımlama hatalarının önüne geçecektir. Soru türü alanının altında kontrol değeri ve puanlama alanları bulunmaktadır. Kontrol alanı bir klasik cevabın doğru kabul edilebilmesi için doğru cevaba ne kadar benzemesi gerektiğini belirlemektedir. Puan alanı, soru eğer doğru ise genel puana eklenecek olan değeri belirlemektedir. Cevap anahtarı oluşturması tamamlandıktan sonra ‘Save’ düğmesine basılarak bir sonraki aşamaya geçilir. Bu aşamada girdi klasöründeki sınav kağıtlarının tamamı otomatik olarak değerlendirilmeye alınacaktır. Değerlendirme işlemi tamamlandıktan sonra sonuçlar bir dosyaya kaydedilir ve kullanıcıya değerlendirme raporu görüntülenir.

4. DENEYSEL ÇALIŞMA

Çalışmanın başarımını değerlendirmek için örnek bir sınav kâğıdı belgesi hazırlanarak önerilen WEB-tabanlı sınav değerlendirme ara-yüzü tarafından test edilmiştir. Şekil 11'de klasik, çoktan seçmeli ve boşluk doldurma şeklinde farklı soru türlerinden örnek soruların yer aldığı ve cevaplandırılmış sınav kâğıdı yer almaktadır. Önerilen çalışmada, ikinci bölümde ele alınmış olan tüm aşamalar otomatik olarak sistem tarafından gerçekleştirilerek, Json dosyası oluşturulmaktadır. Şekil 12'de sunulduğu şekilde, bu Json dosyası daha sonra web tabanlı ara-yüz tarafından rapor halinde sunulmaktadır.



Şekil 11. Test için kullanılan sınav dokümanı.



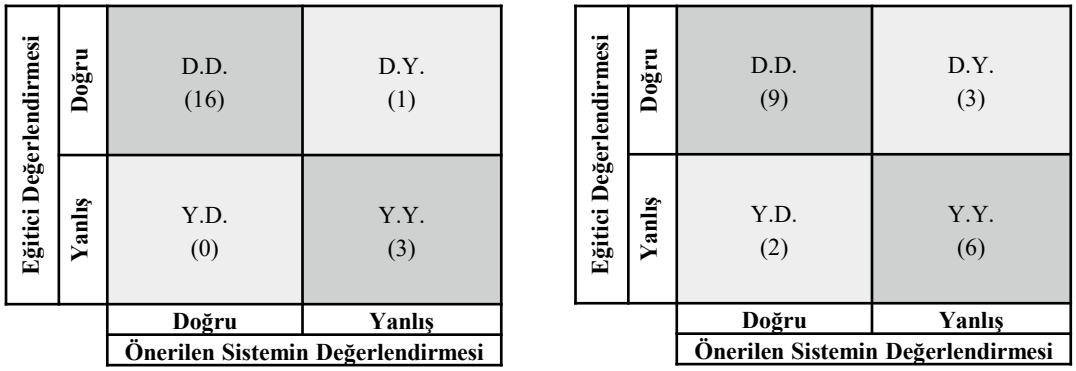
Şekil 12. Değerlendirme sonrasında elde edilen rapor.

Deneysel çalışmalar sonucunda, sadece bir sınav kâğıdının değerlendirilmesi saniyeler içerisinde gerçekleştirilirken sınavda yer alan tüm kâğıtların toplu bir şekilde değerlendirilmesi ise değerlendirilen kâğıt başına yaklaşık 4 saniyede gerçekleştirilmektedir. Tablo 1'de, bir öğrenciye ait sınav kâğıdı için değerlendirme sonuçları yer almaktadır. Görüldüğü üzere, tüm soru tipleri için önerilen sistem tarafında otomatik olarak değerlendirilmiş olup bu örnekte başarım oranı %100 çıkmıştır. Önerilen çalışmanın, kısa cevaplar içeren klasik sorularda başarım oranının çok yüksek olduğu görülmüştür. Ancak, çok uzun cevaplar içeren klasik sorularda değerlendirmenin başarım oranı düşmektedir. Yapılan deneysel çalışmalarda, sınavlarda yer alan çoktan seçmeli soruların yüksek oranda doğru bir şekilde tespit edilerek değerlendirilmesi başarıyla gerçekleştirilmiştir. Önerilen sistemin çoktan seçmeli ve klasik soru türlerindeki değerlendirme başarımını gözlemlemek için, 20 öğrenci için sırasıyla Soru-1 ve Soru-2 için analizler gerçekleştirilmiştir. Bu doğrultuda, Şekil 13'te yer alan Soru-1 ve Soru-2 karmaşıklık matrisleri oluşturulmuştur. Görüldüğü üzere çoktan seçmeli soru tipinde olan Soru-1 için sistemin başarım oranı %95 olarak hesaplanırken, klasik soru tipindeki Soru-2 için bu oran % 75 çıkmaktadır.

Tablo 1

Bir öğrenciye ait sınav kâğıdına için değerlendirme sonuçları.

	Soru 1	Soru 2	Soru 3	Soru 4
Türü	Çoktan Seçmeli	Klasik	Boşluk Doldurma	Çoktan Seçmeli
Cevap	-	"1985'te Bjarne Stroutrup tarafından"	"int" "cout" "i"	B
Benzerlik	-	0,8	1,0	-
Sonuç	Yanlış	Doğru (>=0,7)	Doğru (>=0,7)	Doğru
Başarım	% 100	% 100	% 100	% 100



Şekil 13. Soru-1 (a) ve Soru-2 (b) için karmaşıklık matrisi.

Ayrıca, Şekil 13'teki soru 1 ve soru 2 için verilen karmaşıklık matrisleri kullanılarak denklem 3-8 arasında verilen doğruluk (Accuracy, A), kesinlik (Precision, P), duyarlılık (Recall, R), F1 skoru ve kappa (κ) gibi performans metrikleri hesaplanarak Tablo 2'de listelenmiştir.

$$A = \frac{TP + TN}{TP + TN + FN + FP} \tag{3}$$

$$P = \frac{TP}{TP + FP} \tag{4}$$

$$R = \frac{TP}{TP + FN} \tag{5}$$

$$F1 = 2 \times \frac{P \times R}{P + R} \tag{6}$$

$$\kappa = \frac{A - Pe}{1 - Pe} \tag{7}$$

Burada Pe:

$$Pe = \frac{(TP + FP) \times (TP + FN) \times (TN + FP) \times (TN + FN)}{(TP + TN + FP + FN)^2} \tag{8}$$

Tablo 2'de görüldüğü üzere çoktan seçmeli türünde yer alan soru 1 için, %95 Doğruluk, %94 Kesinlik, %100 Duyarlılık, %97 F1 Skoru ve %83 kappa skoru elde edilmiştir. Bununla beraber klasik türündeki soru 2 incelendiğinde %75 Doğruluk, %75 Kesinlik, %82 Duyarlılık, %78 F1 Skoru ve %49 kappa skoru hesaplanmıştır. Önerilen sistemin bu iki farklı türdeki soruları değerlendirme başarımı incelendiğinde çoktan seçmeli soru türlerinde yüksek başarım elde edilirken klasik soru türlerinde ortalama başarı gösterdiği görülmektedir.

Tablo 2

Soru 1 ve Soru 2 için performans analizi

		Metrik Değerleri				
Soru No	**Soru Türü**	**Doğruluk (A)**	**Kesinlik (P)**	**Duyarlılık (R)**	**F1 Puanı**	**Kappa (κ)**
1	Çoktan Seçmeli	0,95	0,94	1,00	0,97	0,83
2	Klasik	0,75	0,75	0,82	0,78	0,49

5. SONUÇLAR

Bu çalışmada sınıf ortamında kâğıt ve kalem ile yapılan geleneksel sınavların değerlendirilmesine için görüntü işleme teknikleri ve önerilen kelime benzerliklerinin ağırlıkları algoritmasını kullanan yeni bir yöntem önerilmiştir. Önerilen yöntem ile eğitimcilerin zaman kaybı yaşamadan, hızlı, etkili ve verimli bir şekilde, klasik, çoktan seçmeli ya da boşluk doldurma gibi farklı türden soruların yer aldığı sınav kâğıtlarının değerlendirebilmesi amaçlanmıştır. Çalışmada, kamera veya belge tarayıcı yardımıyla bilgisayar ortamına aktarılan sınav kâğıtları, öncelikli olarak soru ve cevaplar şeklinde ayrıştırılarak Json dosyasına kaydedilmektedir. Kaydedilen veriler WEB-tabanlı ara-yüz vasıtasıyla değerlendirilerek raporlanmaktır. Önerilen kelime benzerliklerinin ağırlıkları algoritması ile klasik soru değerlendirmesi hızlı ve güvenilir bir şekilde gerçekleştirilebilmektedir Çalışmanın başarımını değerlendirmek için örnek bir sınav kâğıdı belgesi hazırlanarak WEB-tabanlı sınav değerlendirme ara-yüzü tarafından test edilmiştir. Buna göre kısa cevaplar içeren klasik sorularda başarım oranının çok yüksek olduğu görülmüştür. Öte yandan uzun cevaplı klasik sorularda başarım oranı düşmektedir. Bunun temel sebebi olası cevapları ifade eden cümlelerde yer alan kelimelerin sınırlı olmasından kaynaklanmaktadır. Ayrıca, önerilen sistemin çoktan seçmeli ve klasik soru türleri için başarımı 20 adet öğrenci için değerlendirilmiş olup, çoktan seçmeli soru tipi için % 95, klasik soru tipi için ise % 75 olarak hesaplanmıştır. Önerilen çalışma, Modern C++ ve Python yardımıyla yüksek performansı sayesinde mobil cihazlar ve gömülü sistemlerde kullanıma uygundur. Bununla birlikte, hazırlanan uygulama, çapraz platform olduğu için Linux, Windows veya Mac işletim sistemlerinde de çalıştırılabilir. Çalışmada kullanılan yöntemler farklı uygulamalarda çeşitli amaçlarla oluşturulmuş el yazması belgelerin değerlendirilmesi için kullanılabilir. İleriki çalışmalarda, uzun cevaplardan oluşan klasik soruların değerlendirilmesi için yapay sinir ağı modeli kullanan bir yapının geliştirilmesi ve matematiksel ifadelerin değerlendirilebilmesi için bir matematik motorunun tasarlanması planlanmaktadır.

Hakem Değerlendirmesi: Dış bağımsız.
Çıkar Çatışması: Yazarlar çıkar çatışması etmemişlerdir.
Finansal Destek: Yazarlar bu çalışma için finansal destek almadığını beyan etmemişlerdir.
Yazar Katkıları: Çalışma Konsepti/Tasarım- Y.G., S.S., M.H.B.U.; Veri Toplama- Y.G.; Veri Analizi/Yorumlama- Y.G., S.S., M.H.B.U.; Yazı Taslağı- S.S., M.H.B.U.; İçeriğin Eleştirel İncelemesi- Y.G., S.S., M.H.B.U.; Son Onay ve Sorumluluk- Y.G., S.S., M.H.B.U.; Süpervizyon S.S., M.H.B.U.

Peer-review: Externally peer-reviewed.
Conflict of Interest: The authors have no conflict of interest to declare.
Grant Support: The authors declared that this study has received no financial support.
Author Contributions: Conception/Design of Study-Y.G., S.S., M.H.B.U.; Data Acquisition- Y.G.; Data Analysis/Interpretation- Y.G., S.S., M.H.B.U.; Drafting Manuscript- S.S., M.H.B.U.; Critical Revision of Manuscript- Y.G., S.S., M.H.B.U.; Approval and Accountability- Y.G., S.S., M.H.B.U.; Supervision- S.M.

Kaynaklar/References

Bloomfield, A. (2010, October). Evolution of a digital paper exam grading system. In 2010 IEEE Frontiers in Education Conference (FIE) (pp. T1G-1). IEEE.

Bradski, G., & Kaehler, A. (2008). Learning OpenCV: Computer vision with the OpenCV library. " O'Reilly Media, Inc.".

Chi, Z., & Zhang, B. (2018). A sentence similarity estimation method based on improved siamese network. Journal of Intelligent Learning Systems and Applications, 10(4), 121-134.

Clark, E., Celikyilmaz, A., & Smith, N. A. (2019, July). Sentence mover's similarity: Automatic evaluation for multi-sentence texts. In Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics (pp. 2748-2760).

Çelik, A., (2020). Optik karakter tanımada hata yayılım algoritmalarının performans kıyaslaması. Journal of the Institute of Science and Technology, 10(4), 2328-2340.

Çelik, A., (2021). Eğik Karakter Tanıma Başarısını Arttırmak için Yeni Bir Yöntemin Kullanılması. Harran Üniversitesi Mühendislik Dergisi, 6(1), 1-11.

Gao, Y., Jin, L., He, C., & Zhou, G. (2011, September). Handwriting character recognition as a service: A new handwriting recognition system based on cloud computing. In 2011 International Conference on Document Analysis and Recognition (pp. 885-889). IEEE.

Fowler, M., Chen, B., Azad, S., West, M., & Zilles, C. (2021, March). Autograding" Explain in Plain English" questions using NLP. In Proceedings of the 52nd ACM Technical Symposium on Computer Science Education (pp. 1163-1169).

Gomes Rocha F., Rodriguez G., Andrade E.E.F., Guimarães A., Gonçalves V., Sabino R.F. (2021) Supervised Machine Learning for Automatic Assessment of Free-Text Answers. In: Batyrshin I., Gelbukh A., Sidorov G. (eds) Advances in Soft Computing. MICAI 2021. Lecture Notes in Computer Science, vol 13068. Springer, Cham. https://doi.org/10.1007/978-3-030-89820-5_1.

Ha, L., Yaneva, V., Harik, P., Pandian, R., Morales, A., & Clauser, B. (2020). Automated prediction of examinee proficiency from short-answer questions. Proceedings of the 28th International Conference on Computational Linguistics, pages 893–903 Barcelona, Spain (Online), December 8-13, 2020.

Lagler, K., Schindelegger, M., Böhm, J., Krásná, H., & Nilsson, T. (2013). GPT2: Empirical slant delay model for radio space geodetic techniques. Geophysical research letters, 40(6), 1069-1073.

Lee, J. S., & Hsiang, J. (2020). Patent claim generation by fine-tuning OpenAI GPT-2. World Patent Information, 62, 101983.

Liu, T., Ding, W., Wang, Z., Tang, J., Huang, G. Y., & Liu, Z. (2019, June). Automatic short answer grading via multiway attention networks. In International conference on artificial intelligence in education (pp. 169-173). Springer, Cham.

Liu, Z., Xu, Z., Escalera, S., Guyon, I., Junior, J. C. J., Madadi, M., ... & Tu, W. W. (2020). Towards automated computer vision: analysis of the AutoCV challenges 2019. Pattern Recognition Letters, 135, 196-203.

Llamas-Nistal, M., Fernández-Iglesias, M. J., González-Tato, J., & Mikic-Fonte, F. A. (2013). Blended e-assessment: Migrating classical exams to the digital world. Computers & Education, 62, 72-87.

Marti, U. V., & Bunke, H. (2002). The IAM-database: an English sentence database for offline handwriting recognition. International Journal on Document Analysis and Recognition, 5(1), 39-46.

McGill, T. J., & Klobas, J. E. (2009). A task–technology fit view of learning management system impact. Computers & Education, 52(2), 496-508.

Németh, B., & Tejfel, M. (2016, May). markfactory: Translation-based automatic exam evaluation for mass education. Proceedings of the 11th Joint Conference on Mathematics and Computer Science, Eger, Hungary

Plamondon, R., & Srihari, S. N. (2000). Online and off-line handwriting recognition: a comprehensive survey. IEEE Transactions on pattern analysis and machine intelligence, 22(1), 63-84.

Plötz, T., & Fink, G. A. (2009). Markov models for offline handwriting recognition: a survey. International Journal on Document Analysis and Recognition (IJDAR), 12(4), 269.

Rajala, T., Kaila, E., Lindén, R., Kurvinen, E., Lokkila, E., Laakso, M. J., & Salakoski, T. (2016, February). Automatically assessed electronic exams in programming courses. In Proceedings of the Australasian computer science week multiconference (pp. 1-8).

Sanuvala, G., & Fatima, S. S. (2021, February). A Study of Automated Evaluation of Student's Examination Paper using Machine Learning Techniques. In 2021 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS) (pp. 1049-1054).

Shaukat, Z., Ali, S., Xiao, C., Sahiba, S., & Ditta, A. (2020). Cloud-based efficient scheme for handwritten digit recognition. Multimedia Tools and Applications, 79(39), 29537-29549.

Smith, R. (2007, September). An overview of the Tesseract OCR engine. In Ninth international conference on document analysis and recognition (ICDAR 2007) (Vol. 2, pp. 629-633). IEEE.

Solak, S., Altınışık, U., Yıldız, U., & İnal, M. (2016). Örgün Öğretim Derslerinin Moodle Öğrenme Yönetim Sistemi Kullanılarak Sunulması Deneyimi. Eğitim ve Öğretim Araştırmaları Dergisi, 5(2), 348-360.

Solak, S., Ucar, M. H., & Albadwieh, M. (2020). Computer-based evaluation to assess students' learning for the multiple-choice question–based exams: CBE-MCQs software tool. Computer Applications in Engineering Education, 28(6), 1406-1420.

Şahin, İ., Uçar, M. H. B., & Solak, S. (2022). Otomatik Türkçe Kartvizit Tanıma için Bulut Tabanlı WEB Uygulama Tasarımı ve Performans Değerlendirmesi. Gazi Üniversitesi Fen Bilimleri Dergisi Part C: Tasarım ve Teknoloji, 10(1), 118-134.

Tashu, T. M., Esclamado, J. P., & Horvath, T. (2019, June). Intelligent on-line exam management and evaluation system. In International Conference on Intelligent Tutoring Systems (pp. 105-111). Springer, Cham.

Tomić, S., Paunović, V., & Bosnić, I. Computer-based question and exam evaluation in summative knowledge assessment. In 2020 43rd International Convention on Information, Communication and Electronic Technology (MIPRO) (pp. 1520-1525). IEEE.

Villalon, J. (2012, July). An eMarking tool for paper based evaluations. In 2012 IEEE 12th International Conference on Advanced Learning Technologies (pp. 43-45). IEEE.

Vij, S., Tayal, D., & Jain, A. (2020). A machine learning approach for automated evaluation of short answers using text similarity based on WordNet graphs. Wireless Personal Communications, 111(2), 1271-1282.

Walvoord, B. E., & Anderson, V. J. (2011). Effective grading: A tool for learning and assessment in college. John Wiley & Sons.

Yorke, M. (2009). Faulty signals? Inadequacies of grading systems and a possible response. In Assessment, learning and judgement in higher education (pp. 1-20). Springer, Dordrecht.