

# Müzik Prodüksiyonunda Akıllı Sistemler ve Bir Otomatik Ses Miksaj Hazırlığı Uygulaması Modeli

Arş. Gör. İsmet Emre YÜCEL  
(ITU, MIAM-Müzik Doktora Programı)  
Sakarya Üniversitesi  
Devlet Konservatuarı  
Sakarya – Türkiye  
[eyucel@sakarya.edu.tr](mailto:eyucel@sakarya.edu.tr)

Doç. Dr. Taylan ÖZDEMİR  
İstanbul Teknik Üniversitesi  
Türk Müziği Devlet Konservatuarı  
Müzik Teknolojisi Bölümü  
İstanbul – Türkiye  
[ozdemirtaylan@itu.edu.tr](mailto:ozdemirtaylan@itu.edu.tr)

**Öz:** Sayısal sistemlerin kullanımlarının yaygınlaşması her alanda olduğu gibi müzik üretimi ve tüketiminde de büyük değişimlere neden olmuştur. Günümüzde, endüstri 4.0'ın sağlamış olduğu devrimsel yeniliklerden biri olan akıllı sistemlerin ve bu sistemlerin sağladığı yöntemlerin müzik prodüksiyon aşamalarında karşılaşılan sorunların çözümündeki kullanımları, akıllı müzik prodüksiyonu (Intelligent Music Production) araştırma alanında yer almaktadır. Akıllı müzik prodüksiyonu çalışmaları, otomatik ses miksağı, semantik ve algısal açıdan audio sinyal işlemeyi kolaylaştırmayı hedefleyen arayüzler, audio sistemlerde ölçme (metering) ve tanılama (diagnosis), interaktif ortamlarda yer alan sanal gerçeklik (virtual reality) tabanlı uygulamalar ve video oyunlarındaki ses işleminin otomatikleştirilmesi gibi geniş bir çalışma alanını kapsamaktadır.

Bu çalışmada müzik prodüksiyon aşamalarındaki ihtiyaçlar ses mühendisliği perspektifinden değerlendirilecek ve bu ihtiyaçların giderilmesinde makine öğrenmesi tabanlı akıllı uygulamaların yeri tartışılacaktır. Bu bağlamda geliştirilebilecek uygulamalara örnek olarak çok kanallı ses miksağı hazırlığı için audio kanal alt-gruplamalarını (subgrouping) otomatik olarak yapabilen makine öğrenmesi temelli bir yazılım tanıtılacaktır.

**Anahtar Sözcükler:** Müzik Teknolojileri, Akıllı Müzik Prodüksiyonu, Ses Miksağı ve Hazırlığı, Makine Öğrenmesi, Akıllı Sistemler, Müzik Yazılımları.

## Intelligent Systems in Music Production and an Application of Mix-Preparation Software Model

**Abstract:** The widespread usage of digital systems has also caused changes in music production and consumption similar to the developmental phase in many fields. Today, the utilization of intelligent systems which are considered to be prominent evolutionary innovations in industry 4.0, and their beneficial methods for solving problems in music production stages are presented in the research field of Intelligent Music Production (IMP). This research field covers a broad scope of studies in which automatic mixing, signal processing with semantic and cognitive audio interfaces, metering and diagnostic solutions, interactive virtual reality-based applications and automatic signal processing in video games take place.

In this study, the applicability of machine learning-based intelligent systems in meeting the needs and requirements of music production steps will be discussed in terms of sound engineering. Herein, a machine learning-based automatic subgrouping software for multitrack audio mix-preparation will be introduced as an example of an IMP solution.

**Keywords:** Music Technology, Intelligent Music Production, Audio Mixing and Mix Preparation, Machine Learning, Intelligent Systems, Music Software.

### 1. Giriş

Müzik prodüksiyonu ses mühendisliği perspektifinden bakıldığında kayıt, miksaj ve kalıplama (recording, mixing, mastering) olacak şekilde üç ana aşamada ele alınmaktadır. Bu aşamalar birbirleri ile doğrudan ilişkili olup, bu ilişki sonuçta ortaya çıkacak olan müzikal eserin ticari başarısı ile doğrudan ilişkilidir. Akıllı sistemlerin müzik prodüksiyonu ve ilgili alt aşamalarında karşılaşılan teknik ve artistik zorlukların çözümlerinde kullanılmaları ise akıllı müzik prodüksiyonu araştırma alanının ortaya çıkmasını sağlamıştır. Bu alanda yapılan çalışmalar ile müzik üretimini daha hızlı ve kolay bir hale gelecektir. Müzik prodüksiyonundaki alt aşamalardan biri olan çok kanallı miksağı hazırlığı, ses mühendisleri arasında oldukça yaygın bir uygulama olup, özellikle proje içerisindeki audio kanal sayısı arttıkça bu kanalların ait oldukları çalgı ailelerine ve aranjman içerisindeki fonksiyonlarına göre karmaşık bir hale gelmektedir. Bu çalışmada müzik prodüksiyonunda akıllı sistemler perspektifinden bakarak, çok kanallı miksağı hazırlığında alt-gruplamaları makine öğrenmesi modeli ile otomatik hale getirmeyi hedefleyen bir yazılım prototipi tanıtılacaktır. Yazılımın çalgı ailelerine göre otomatik alt-gruplama özelliği 80 farklı çok kanallı proje ile toplamda 1428 ses dosyası ile test edilmiştir. Ses dosyalarına verilen isimler ve makine öğrenmesi modeli ile elde edilen sonuçlar kıyaslandığında genel başarının 62.5% olduğu gözlemlenmektedir.

## 1.1. Müzik Prodüksiyonu

Müzik prodüksiyonu, bir ses kayıt stüdyosu içinde ya da dışında müziğin yaratıcı ve artistik açıdan geliştirilmesini ifade eden bir terim olmasının yanı sıra, müziğin kitlelere fiziksel ya da çevrim içi şekilde bir ürün olarak ulaştırıldığı geniş bir süreçtir [1]. Müzik üretimi, prodüksiyon öncesi, prodüksiyon ve prodüksiyon sonrası olacak şekilde 3 temel gruba ayrılmaktadır. Besteleme, düzenleme, prodüksiyon öncesini süreçleri; kayıt, miksaj, kalıplama prodüksiyon sürecini; müziğin kitlelere ulaştırıldığı dağıtım ile ilgili işlemler ise post prodüksiyon süreçlerini kapsamaktadır.

Tarihsel sürece bakıldığında müzik prodüksiyonunun gelişiminin kayıt teknolojilerinin gelişimi ile paralel bir seyirde olduğu görülmektedir. 19. YY başlarında bir ses çatalı üzerinde oluşan titreşimleri is kaplı bir silindire kaydeden Vibrograf (Thomas Young), yine sonraki yıllarda benzer bir düzenek olan vibraskop (Jean-Marie Constant Duhamel) sesin oluşumunu incelemek amacıyla üretilmiş, daha çok osilaskopa benzeyen ilkel cihazlardı. 1860 yılında Edouard-Leon Scott de Martinville tarafından tarihteki bilinen ilk ses kaydını gerçekleştirdiği geri çalma (playback) özelliği olmayan fonografı icat etti. Ardından, Edison'un 1877'de geri çalma yapabilen ilk kayıt cihazı olan fonografı (ses-yazar, sound-writer) icat etti. 1888'de gramafon'un Emile Berliner tarafından icat edilmesi ve ticari potansiyelinin vurgulanması, müzik prodüksiyonunda günümüzde de kullanılan (records, master disc gibi) temel terminolojinin de ortaya çıkmasına katkı sağlamasının yanı sıra kayıt endüstrisinin ortaya çıkmasına da neden olmuştur. 20. YY gelindiğinde ise, kondenser mikrofonun icadı, vakum tüp'ün (Audion) Lee de Forest tarafından müziğin kablosuz aktarılması amacıyla icadı audio teknolojilerindeki önemli gelişmelerdir 1925'lere kadar olan ve akustik ya da salt-mekanik kayıt teknolojisinin yer aldığı bu dönemin ardından silindir diskin üzerine elektronik olarak yazma girişimleri başladı [2]. İlk çalışmaları 1920'li yıllarda Almanya'da başlayan, ardından 1945 yılında Amerika'daki tanıtımından sonra, analog teyp kaydediciler profesyonel ve kişisel kayıt stüdyolardaki kullanımları yaygınlaştırmıştır [3]. 1960'lı yıllardan önce orkestra kayıtları büyük odalarda bir ya da birkaç mikrofon ile yapılmakta ve odanın akustik özellikleri akustik özellikleri kaydın ayrılmaz bir parçası olarak karışımıza çıkmaktaydı. 1960'larda 4-8 kanallı teyp kaydedicilerin ortaya çıkması ile çok kanallı ses kayıt teknolojileri bir standart haline geldi. 1970'lere gelindiğinde, 24 kanallı kaydediciler yaygınlaşması ile stüdyo kayıtları daha izole hale gelmeye başladı bu nedenle harici yankıma (reverb) ve eko odaları (echo chambers) gerekli hale geldi. 1980'lerde sayısal yankıma işlemcileri geliştirilmiş ve geniş bir kullanım alanı bulmuşlardır. Yine aynı dönem stüdyolardaki kayıt sistemlerinde kanal sayıları 72'ye kadar çıkmaktaydı. 1990'larda sayısal teknolojilerin ses kayıt stüdyolarına girmesi ile 24-32-48 kanallı sayısal kayıt sistemleri yaygınlaştı. 2000'li yıllara gelindiğinde ise müzik prodüksiyonu "in the box" olarak nitelendirilen, tamamen bilgisayar temelli ses iş istasyonları (SSİİ) üzerinden gerçekleştirilmeye başlandı [4]. Günümüzde tüm müzik prodüksiyon aşamaları SSİİ (DAW-Digital Audio Workstation) üzerinde gerçekleştirilmektedir.

Müzik endüstrisindeki değişimleri salt teknolojik gelişmelerden ziyade, bu teknolojik gelişmelerin müzik üretiminde yarattığı paradigma değişimlerini de değerlendirmek gerekir. Bu bağlamda, gramofon döneminde müziğin kaydedildiği ilk mekânlar konser salonları ve sahneler iken, manyetik bant teknolojisi ile kayıt ve dinleme odası gibi amaca yönelik mekân ihtiyaçları ortaya çıkmış ve ses stüdyoları doğmuştur. Çok kanallı kayıt döneminde manyetik teypler için üstüne okuma (over-dubbing) teknolojisinin yanı sıra stereo kayıt oynatabilen teknolojinin kullanıcılar arasında yaygınlaşması ile çalgıların stereo dinleme alanına (psiko-akustik açıdan) nasıl yerleştirileceği hakkında hem kayıt (mikrofonlama teknikler gibi) hem de miksaj ile ilgili yaklaşımlar ve yöntemler geliştirilmiştir. Sayısal kayıt sistemleri, müzik kaydı üzerinde aslına-zarar-vermeyen (non-destructive) düzenleme yöntemlerine izin vermeleri açısından büyük avantajları sağlamıştır, böylece gramofon teknolojisinde imkânsız, bant teknolojilerinde ise oldukça zor olan "audio editing" oldukça kolay hale gelmiştir. Günümüzde SSİİ'ler artık taşınabilir teknolojilerin içerisinde dahi yer almaları sayesinde müzik üretimi son kullanıcılar (yani müziği tüketen kitleler) arasında yaygınlaşmıştır. Çevrim içi servislerin (streaming) ve özellikle müzik paylaşım platformlarının yaygınlaşması ile müzik prodüksiyonun giderek şirketlerin tekeline kurtulduğunu söyleyebiliriz. İşte tam bu noktada akıllı müzik prodüksiyonu alanında yapılan çalışmaların müzik endüstrisinde profesyonellerin yanı sıra yarı-profesyonel ve amatörler için sunduğu/sunacağı avantajlar aşikârdır.

## 1.2. Prodüksiyon aşamaları

Müzik prodüksiyon aşamaları kayıt, miksaj ve kalıplama (mastering) olarak adlandırılan üç ana aşamadan oluşmaktadır. Bu aşamalar sadece mühendisliğin (elektrik-elektronik ve haberleşme gibi) ve mimarlığın da (kapalı mekân akustiği) dahil edilebileceği farklı disiplinlerdeki teknik konular ve literatür ile derinden ilişkili değil, aynı zamanda bu alanlarda var olan bilgi birikiminin en etkili biçimde kullanılarak artistik kararların alındığı aşamalardır.

Ses kaydı kaynaktan çıkan akustik enerjinin bir dönüştürücü (transducer) vasıtası ile elektrik sinyallerine çevrilip ortama (medium) yazılması olarak basitçe tanımlanabilir. Ancak, kaydın yapıldığı mekânın akustik özellikleri, mikrofon tipi ve yerleşimi, bununla birlikte kayıt zincirinde yer alan ve kayıt sinyalinin üzerinden analog olarak geçtiği tüm donanımlar kaydın teknik ve artistik niteliğini doğrudan etkilemektedir.

Akustik ve psiko-akustik ses mühendisliğinin sıklıkla referans aldığı iki alandır. Ses kayıt stüdyolarındaki kayıt ve kontrol odaları genelde küçük mekanlardır, bu nedenle Kayıt ve kontrol odalarındaki hacim, şekil ve form gibi mimari özellikler oda içerisindeki yankıma süresini, yansıma, eko, oda modları vb. akustik özellikleri değiştirmektedir. Profesyonel ve yarı profesyonel kayıt stüdyolarında yalıtım ihtiyaçlarının giderilmesi için (yüzer sistem tercihleri, vokal ve davul kabin tasarımları, duvarları ses geçiş hesapları gibi) mimari akustik literatürü önemli bir rol oynamaktadır. Öyle ki, akustik özelliklerinin kontrolü açısından, ses kayıt stüdyolarının geometrik özelliklerinin standartlaştırıldığı geniş diyebileceğimiz bir oda tasarımı literatürü birikmiştir. DDA (Directional Dual Acoustics), LEDE (Live-End, Dead-End), Non-Environment, Toyoshima odaları geometrik açıdan kontrollü kontrol odası tasarımlarıdır[5]. Yine psiko-akustiğin sağlamış olduğu literatürde yer alan örneğin; kaynağın lokalizasyonu (iki kulak arasındaki zaman ve yoğunluk farkı açısından), genlik ve frekans algısı, genlik ve frekans algısı arasındaki ilişkiyi gösteren eşit gürlük çizelgesi (equal-loudness contour), frekans bant genişliği ve maskeleyme (frequency masking), kayıp temel ses (missing fundamental), duysal hafızanın lokalizasyona etkisi ile ilgili olan Fransen etkisi, ses ve o sesin tekrarının bir bütün ya da gecikme olarak algılanıp algılanmadığını açıklayan Haas etkisi gibi çalışmalar sesin kaydı ve işlenmesi sırasında baş vurulan temel bilgiler olarak karşımıza çıkmaktadır [6].

Kayıt sırasında ise mikrofon seçimi ve yerleşimi müzik prodüksiyonu tipine (pop, caz, orkestra vb.) göre değişiklikler gösterir. Burada, örneğin bir mikrofonun kazanç (gain) seviyesinin ayarlanması teknik bir konu iken, mikrofonların kapsül tipleri, yerleşimleri ve kaynağa olan uzaklıkları, kayıt aşamalarındaki müzik prodüksiyon tiplerine bağlı artistik ve estetik tercihlere tipik örneklerdir ve telekomünikasyondaki sistemler için geçerli standartlar ses mühendisliğinde alanında da referans olarak kabul görmektedirler.

Müzik prodüksiyonunda bir müzik parçasını oluşturan birden çok audio kanalın kaydından sonra bu kanalların karıştırma aşaması gelir. Çok kanallı ses miksajı (multitrack mixing) olarak adlandırılan bu aşamada, şarkıda yer alan tüm audio kanalların dinleme alanındaki pozisyonları müzik içerisindeki fonksiyonları ve önemleri açısından değerlendirilerek belirlenir. Dinleme alanındaki bu yerleştirme süreci yine teknik ve artistik kararların alındığı karmaşık bir süreçtir. Alınan kararlar müzik tarzı, çalgı sayısı, çalgıların önemi ve dahası şarkının duygu durumuna göre şekillenirken, bu kararların uygulanması ise genel anlamda zaman temelli, frekans temelli ve genlik temelli olacak şekilde sınıflandırılan ses işlemcileri ve bu işlemcilerin kombine edildiği diğer işlemciler ile gerçekleştirilir (Tablo 1).

**Tablo 1:** Audio işlemciler

Frekans Temelli	Genlik Temelli	Zaman Temelli	Kombine işlemciler
Ekolayzırlar	Kompresörler		
Dinamik Ekolayzırlar	Dinamik kompresörler	Yankıma (reverb)	Vibrato
Filtreler	Kapılar (Gates)	Geçikme (delay)	Flanger
Distorsiyon (Distortion)	Sınırlayıcılar (Limiters)		Phaser
Harmonik ekleyiciler (Harmonizers)	Genişleticiler (Expanders)		Tremolo
	Geçiş düzenleyiciler (Transient designers)		

Çok kanallı ses miksajına başlamadan önce miksaj hazırlığı (mix preparation) özellikle büyük projeler için bir alt uygulama olarak sıklıkla karşımıza çıkmaktadır. Miksaj hazırlığı miksaja başlamadan önce kanalların organize edilmesi, gruplandırılması, gerekli ise tekrardan isimlendirilmesi, kanal seviyelerinin hızlıca ayarlandığı kaba miksajın (rough mix) gerçekleştirildiği bir aşamadır. Owsinsky [7] miks hazırlığının miksaj sırasında karşılaşılabilecek muhtemel hataların ve çabanın minimize edilmesi açısından önemli olduğunu vurgulamıştır (s. 21).

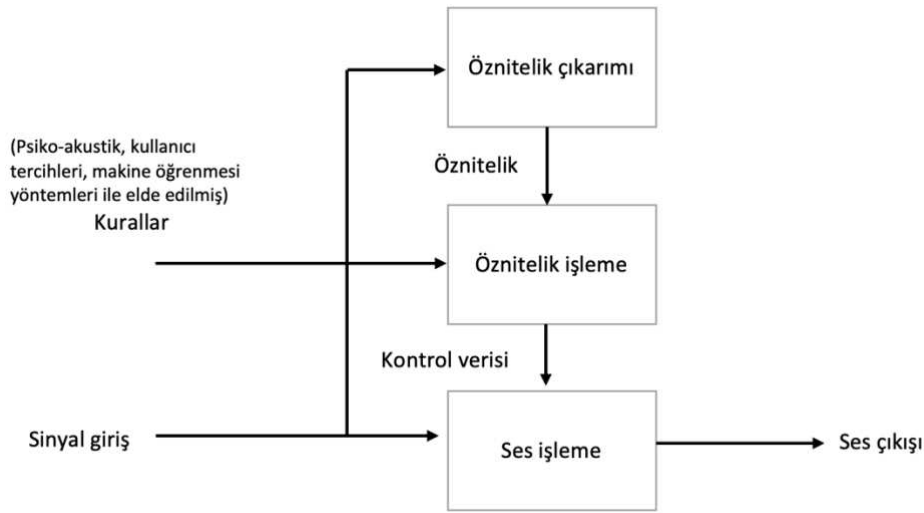
Çok kanallı miksaj aşamasını kalıplama aşaması takip eder. Bu aşamada bir müzik albümdeki her bir eserin duysal açıdan eşit seviye ve karakterde olması amacı ile son dokunuşlar yapılır. Bu noktada şarkıların çözünürlüğü ile ilgili örnekleme oranı dönüşümü, bit derinliği ilgili işlemler yapılmakta ve şarkıların gürlük seviyeleri yayınlanacakları ortama (örneğin CD, çevrimiçi servisler, radyo, TV vb.) göre ayarlanmaktadır. Ses mühendisleri Uluslararası Telekomünikasyon Birliği (ITU) ve Avrupa Yayıncılık Birliği (EBU) tarafından önerilen gürlük seviyelerini referans almaktadırlar.

### 1.3. Akıllı Müzik Prodüksiyonu

Akıllı müzik prodüksiyonu sistemleri prodüksiyonun her aşaması için ses mühendisinin en az müdahalesi ile gerekli audio işlemeyi otomatik olarak gerçekleştirmeyi hedefleyen sistemlerdir. Literatürde audio uygulamaları ile ilgili ilk otomatik sistem

canlı seslendirmelerde geri besleme problemini çözmek amacı ile otomatik kazanç kontrolü sağlayan bir devre tasarımıdır[8]. Sayısal sistemlerin sağladığı avantajlar ile kazanç kontrolünün yanı sıra stereo panlamayı, frekans dengelemeyi, gecikme ve polarizasyon düzeltmeyi otomatik gerçekleştirmeyi hedefleyen yöntemler ortaya atılmaya başlandı. 2007 ve sonrasında ses miksajını otomatikleştirmeyi hedefleyen yeni yöntemler akıllı müzik prodüksiyonu alanındaki araştırmalarda önemli bir yere sahiptir. De Man [9] otomatik miksaj sistemlerini dört ana kategoride sıralamıştır: algoritmanın giriş sinyaline göre, dışarıdan gözlemlemeye izin vermeden çıkış verisini ürettiği kara kutu sistemler, ses mühendisine yardımcı asistan sistemler, semantik ara yüzler, audio sinyali görsel açıdan daha iyi inceleye imkân sağlayan ölçme ve tanılama sistemleri.

Akıllı müzik prodüksiyonu temel araçları hiç kuşkusuz ki otonom ses işlemcileridir. Ses işlemcileri otonom hale getirebilmek işi ses mühendisinin audio işlemeye olan yaklaşımını makine öğrenmesi, psiko-akustik kurallar ile modellemek ve audio işlemcinin bu kurallara göre giriş sinyalini değerlendirmesi sağlamaktır (Şekil 1) [10]. Reiss [11] çok kanallı audio sinyalinin işlenmesinde side-chain sinyalinin birden fazla kanalla sürekli değişiklik gösterdiğini belirtmiş, buna çözüm olarak da adaptif sayısal audio efektlerin (cross-adaptive digital audio effects) kullanımı önermiştir.



Şekil 1: Akıllı ses işleme sistemleri için akış diyagramı

Öznitelik çıkarma çalışmaları audio içerik analizi (audio content analysis) araştırma alanı içerisinde yer almaktadır. Aslen daha geniş bir kapsama sahip olan müzikten bilgi edinme çalışmalarının (MIR-Music Information Retrieval) önemli bir parçasını oluşturan audio içerik analizinde, sinyal işleme yöntemleri ile audio sinyalden daha anlamlı ve yorumlanabilir veriler çıkartılır. Öznitelikler genel olarak düşük seviye (low-level) ve olarak kategorize edilirler. Düşük seviye öznitelikler insan tarafından doğrudan yorumlanamazken, yüksek seviye öznitelikler müzikal açıdan örneğin; tempo, müzikal yapı vb. insan tarafından yorumlanabilen sonuçlar olarak karşımıza çıkar [12].

Giannakopoulos ve Pikrakis [13] öznitelikleri zaman temelli (temporal) ve izgesel (spectral) olmak üzere ikiye ayırmıştır (Tablo 2). Kısa zaman (tipik olarak 2-4 mili saniye) çerçevelerinde (time-window) hesaplanan ve düşük seviye olarak da ifade edilen bu özniteliklerin daha uzun çerçevelerdeki istatistiklerinin yorumlanması ile audio sinyal hakkında bilgi edinimini sağlar.

<b>Tablo 2 Audio Öznitelikler (audio features)</b>	
<b>Zaman (Temporal)</b>	<b>Enerji (Energy)</b> <b>Sıfır Geçiş oranı (Zero-Crossing Rate)</b> <b>Enerji Düzensizliği (Entropy of Energy)</b>
<b>İzgesel (Spectral)</b>	<b>İzgesel Merkez (Spectral Centroid)</b> <b>İzgesel Yayılma (Spectral Spread)</b> <b>İzgesel Değişim (Spectral Flux)</b> <b>İzgesel Etek (Spectral Roll-off)</b> <b>MFCCs (Mel Frequency Cepstrum Coefficients)</b> <b>Doğunluk vektörü (Chroma Vector)</b>

Çıkarılan öznitelikler  $M \times N$  matris olacak şekilde sıraya konurlar. Matrisin son sütunu ise sınıf hedefleri (class targets) belirtir. Matrisin satır ve sütun büyüklükleri sınıflar içerisindeki ses dosyalarının sayısı, her birinin uzunluğu, çıkartılan özniteliklerin sayısı ve bu öznitelikler üzerinde yapılan istatistiklere göre değişiklik gösterir. Sonuçta oluşan bu matris bir veri setidir ve makine öğrenmesi algoritmasının eğitim (training) aşamasında ve test aşamasında kullanılır.

## 2. Çok Kanallı Miksaj Hazırlığı Prototipi

Çok kanallı ses miksaj hazırlığını otomatikleştirmeyi amaçlayan bir yazılım için ilk olarak bu yazılımın uyumlu bir şekilde çalışabileceği, tercihen profesyonel SSİİ belirlendi. Hedeflenen yazılımın temel kabiliyeti, çok kanallı miksaj oturumunda gerekli olan alt-gruplama işlemi otomatik gerçekleştirmek olduğu için, tercih edilecek SSİİ; kanal yaratma, çok kanallı projelerin içerisindeki ses dosyalarını içeri aktarabilme ve bu ses dosyalarını amaca yönelik olarak sıralayabilmeye imkan sağlamalıdır.

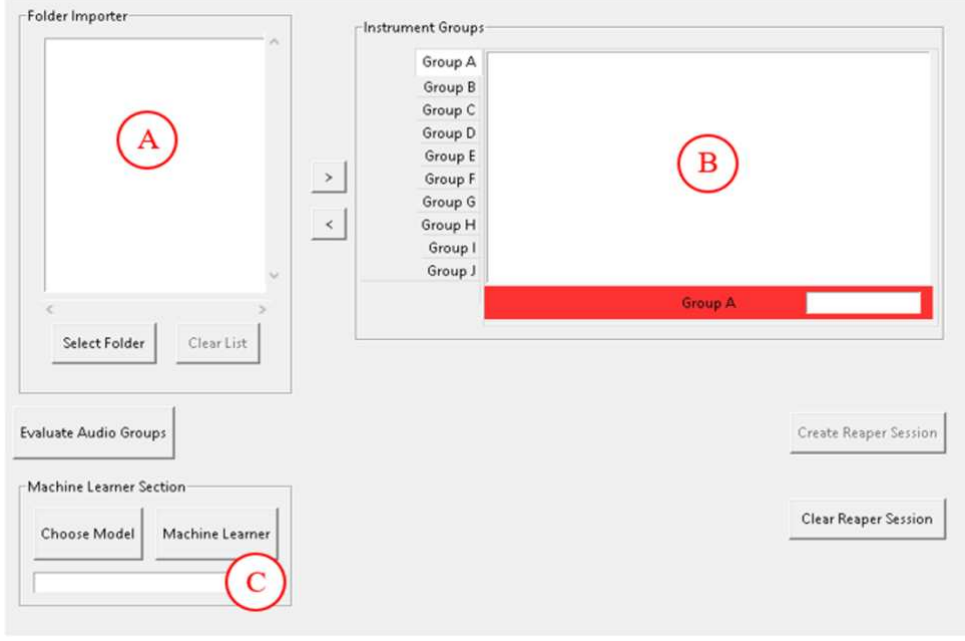
Yazılımın uygulanacağı SSİİ belirlendikten sonra, çok kanallı ses miksaj hazırlığından beklenen; alt-gruplama, alt-grupların isimlendirilmesi ve renklendirilmesi, sonra olarak da bu alt-grupların istenilen amaca yönelik olacak şekilde düzenlemesi gibi temel görevleri yapabilecek yazılımın geliştirilmesi için bir yazılım geliştirme platformu belirlenmiş ve bir grafik ara yüz tasarlanmıştır.

Hedeflenen yazılımın belirtilen temel ses miksaj hazırlığı görevini yerine getirebilmesi için, bir güdümlü makine öğrenmesi (supervised machine learning) modelinin kullanılması uygun görülmüştür. Bu modelin oluşturulmasında kullanılacak veri seti için tek kaynaklı (single-sourced) ve tekrarlı ses dosyalarından (audio-loops) oluşan bir kütüphane tercih edilmiş, bu kütüphanenin içeriğindeki ses dosyaları çalgı ailelerini temsil edecek şekilde gruplandırılmıştır. Böylece alt-grupların çalgı gruplarına uygun şekilde gruplandırılmaları hedeflenmiştir. Yazılımın performansı gerçek çok kanallı projeler ile test edilmiştir.

### 2.1. SSİİ (Sayısal Ses İş İstasyonu) Seçimi ve Yazılım Geliştirme Platformu Tercihi

Piyasada birçok profesyonel SSİİ bulunmaktadır; ancak bunlar arasında hem açık kaynak hem de prodüksiyon ihtiyaçlarını karşılayacak profesyonel bir yazılım olması nedeniyle Reaper en iyi seçenek olarak karşımıza çıkmaktadır. ReaScript adı verilen ara yüzü ile, EEL2, Lua, ve Python yazılım dilleri ile entegrasyonu olan Reaper, açık kaynak kodlu, geliştirilmeye açık, tam profesyonel bir SSİİ'dir [1]. Çok kanallı miksaj hazırlığı yazılım prototipinin ara yüzü ve çalgı tanımlama görevini gerçekleştirecek makine öğrenmesi uygulaması için Python yazılım geliştirme platformu tercih edilmiştir.

Bir Reaper oturumunu uzaktan kontrol edecek şekilde tasarlanan çok kanallı ses miksaj yazılımının ara yüzü Şekil 2'de verilmiştir. Miksaj hazırlığı uygulamasının özelliği, alt gruplamalar ile bu grupların adlandırılma ve renklendirmeleri ile sınırlandırılmıştır. Kullanıcı, tek kaynaklı ses dosyalarını bu ön tanımlı alt gruplara (A'dan J'ye, ön tanımlı 10 adet) atayabilmekte ve her bir grup isim ve renk kodlarını değiştirebilmektedir. Bununla birlikte, bu yazılım alt gruplamaları çalgı sınıflarına göre otomatik olarak yapabilmektedir.



**Şekil 1:** Çok kanallı ses miksaj hazırlığı yazılımı grafik ara yüzü.

Yazılım üç ana pencereden oluşmaktadır. A ile işaretlenen kısım, mevcut çok kanallı projenin içerisinde yer alan audio dosyaların listelendiği pencereyi, B ile işaretlenen kısım Reaper içerisinde oluşturulacak olan 10 adet alt grubun düzenlendiği pencereyi ve C ile işaretli kısım ise kullanıcının veri seti yaratması için ve/veya mevcut veri setlerini kullanmasına olanak sağlayan grafik ara yüz bileşenlerini göstermektedir. Kullanıcı, makine öğrenmesi modelini seçtikten sonra alt grupları oluşturabilir ve ihtiyaca göre otomatik alt gruplama yapabilir. Makine öğrenmesi temelli otomatik gruplamada alt-grup isimleri, makine öğrenmesi modelinde tanımlı çalgı ailelerinin (yani sınıfların) isimlerine bağlı olarak değişmektedir. Kullanıcı alt-grup isimlendirme ve ön tanımlı renklendirmelerine değiştirebilir, ardından projenin miksajı için bir Reaper oturumu yaratabilir.

## 2.2. Kullanılan Ses Kütüphanesi, Dataset ve Deneş

Makine öğrenmesi modelleri bir veri setine ihtiyaç duyarlar. Bu çalışmada, müzik prodüksiyonlarında sıklıkla karşılaştığımız tekrarlı kayıtlardan (audio loops) oluşan bir ticari bir kütüphane tercih edildi. Toplamda 5275 farklı audio dosyası Davul-perküsyonlar, Baslar, Gitarlar, Tuşlular, Yaylılar ve Üflemeliler olacak şekilde 6 çalgı grubu oluşturacak şekilde düzenlendi. Düzenlenen çalgı ailelerinin sahip olduğu audio dosya sayısı Tablo 3'te verilmiştir.

**Tablo 3:** Veri seti içerisindeki sınıf isimleri (çalgı grupları) ve her bir sınıftaki ses dosyası sayıları.

Sınıflar (Çalgı grupları)	Davul ve Perküsyonlar	Baslar	Gitarlar	Tuşlular	Yaylılar	Üflemeliler
Ses Dosyası Sayıları	1965	843	1114	599	395	359

Zaman ve izgesel özniteliklerin ortalamaları, standart sapmaları ve orta değerleri hesaplanarak 5276x106'lık bir veri seti oluşturuldu. Bu veri seti ile SVM (RBF-Radial Basis Function) algoritması eğitilerek bu 6 çalgı grubu için toplamda %94,3 başarı oranı elde edildi. Bu model daha sonra çok kanallı miksaj hazırlık yazılımının çalgı gruplarını otomatik gruplaması amacıyla kullanıldı.

Çalgı gruplaması modeli, 4 farklı müzik tarzında 80 farklı müzik projesinde yer alan toplamda 1428 adet audio dosyası ile test edilmiştir. Sonuç olarak, mevcut makine öğrenmesi modelinin bu dosyalar için belirlediği çalgı ailesi gruplarının, audio

dosyalarının isimleri ile kıyaslandığında %65,2 başarı elde ettiği görülmüştür. Müzik tarzına göre başarı sırası Rock %69.31, Pop %66.84, Caz %66.15 ve Elektronik/Dans %57.97’dir. Çalgı ailelerine göre başarı sırası Davul ve Perküsyonlarda %85.43, Baslarda %67.44, Yaylılarda %61.11, Gitarlarda %54.55, Üflemelelerde %23.40, Tuşlularda ise %15.53 olduğu gözlemlenmiştir (Tablo 4).

**Table 4:** Dört müzik tarzına ve çalgı ailelerine göre başarı oranları.

Müzik Tarzına Göre Başarı Oranları	
Rock	69.31 %
Pop	66.84 %
Caz	66.15 %
Elektronik/Dans	57.97 %

Çalgı Ailelerine Göre Başarı Oranları	
Davul ve Perküsyonlar	85.43%
Baslar	67,44%
Yaylılar	61.11%
Gitarlar	54.55%,
Üflemeleler	23.40%
Tuşlular	15.53%

### 3. Sonuç

Bu çalışmada müzik prodüksiyonunda akıllı sistemler ve bu bağlamda çok kanallı ses miksaj hazırlığını otomatik gerçekleştirebilen bir yazılım prototipi tanıtılmıştır. Tasarlanan yazılımın başarısı 80 farklı müzik prodüksiyon projesi içerisinde 1428 adet tek kaynaklı ses dosyası ile test edilmiştir. Bu deneyde müzik tarzlarında çok büyük değişiklikler olmamakla birlikte, bu çalışmada kullanılan makine öğrenmesi modelinin Davul ve Perküsyon sınıfı için en iyi sınıflandırma başarısını sağladığı görülmektedir (85.43%). Tablo 4’teki sonuçlar değerlendirildiğinde, mevcut veri seti ve bu veri seti ile yaratılan makine öğrenmesi modeli vurmali çalgıların gruplandırılmasında oldukça başarılıdır. Buna karşın tuşlu çalgıların sınıflandırılma başarısı düşüktür (15.53%). Burada düşük başarı oranları için şu iki önemli noktayı irdelemek gerekmektedir. Bunlardan birincisi veri setinin kapsamı, ikincisi ise prodüksiyon içerisinde kullanılan ses dosyalarının isimlerinin içerikleri ile olan uyumsuzluğudur.

Veri seti kapsamı açısından, çalgı ailelerinin kapsam ve genişliğinin amaca yönelik olacak şekilde düzenlenmesi konusu; örneğin gitar ailesi tasarlanırken veri setine hangi gitar tipinin (akustik, elektrik, ukulele, mandolin vb.) dahil edileceği meselesinin yanı sıra bu kayıtlarda olan efektlere (örneğin distorsion, eko, yankılama vb.) bağlı kaynağın tınısını etkileyen durumlardır. Buna göre çalgı aileleri örneğin akustik gitar, efektli gitar vb. tek bir aile grubu yerine daha fazla alt gruba ayrılabilir ve veri seti buna göre tasarlanabilir.

Çalgıların aranjmandaki fonksiyonları ve bu bağlamda isimlendirilmeleri ile ilgili uyumsuzluk durumu ise başarı oranının ölçümü ile ilgili bir problem olarak karşımıza çıkmaktadır. Örneğin özellikle Elektronik/Dans tarzında yer alan ve “synth” olarak adlandırılan ses dosyalarından bazılarının tınısının yaylı çalgıları çağrıştırmakta ve buna paralel çalgı tanımlama modeli bu ses dosyalarını, her ne kadar adına göre tuşlu çalgılar grubundan tanımlaması beklense de yaylı ailesine dahil etmektedir. Yine “synth” olarak adlandırılan ses dosyalarından bazılarının arpejleme gibi ritmik öğeler ile süslendirilmiş olmaları ve/veya katmanlı seslerden üretilmeleri çalgı tanımlama modelinin bu gibi ses dosyalarını Davul ve Perküsyon grubu içerisinde tanımlamaya yatkın olduğu sıklıkla gözlemlenmektedir.

Sonuç olarak tanıtılan yazılım prototipi, akıllı müzik prodüksiyonu çalışma alanının merkezinde yer alan çok kanallı ses miksajının otomatikleştirilmesinde faydalı olacak bir çalışmadır. Bu çalışmada takip edilen metodolojide görüldüğü gibi, çalgı tanımlama algoritmalarının geliştirilmesi ve SSI’lere uygulanması, prodüksiyon aşamalarında zaman alan rutinlerin otomatik olarak gerçekleştirilmesini sağlayacak ve müzik üretiminin daha kolay hale getirecektir.

- [1] R. Hepworth-Sawyer and C. Golding, *What is Music Production?: a producers guide, the role, the people, the process*, Routledge, 2010.
- [2] R. James. Burgess, *The History of Music Production*, Oxford University Press, 2014.
- [3] D. Miles. Huber and R. E. Runstein, *Modern Recording Techniques*, 7th ed., Routledge, 2010.
- [4] H. Weekhout, *Music Production Learn How To Record, Mix and Master Music*, 3rd ed., Routledge, 2019.
- [5] P. R. Newell, *Recording Studio Design*, 4th ed., Routledge, 2017.
- [6] F. Alton Everest and Ken C. Pohlmann, *Master Handbook of Acoustics*, 7th ed., McGrawHill, 2021.
- [7] B. Owsinski, *The Mixing Engineer's Handbook*, 3rd ed., Course Technology PTR, 2014.
- [8] D. Dugan, "Automatic Microphone Mixing", *AES*, vol. 23, s. 442–49, 1975.
- [9] B. de Man, J. D. Reiss, and R. Stables, "Ten Years of Automatic Mixing", *Proceedings of the 3rd Workshop on Intelligent Music Production*, Salford, UK, 2017.
- [10] B. de Man, R. Stables, and J. D. Reiss, *Intelligent Music Production*, Routledge, 2020.
- [11] J. D. Reiss, "Intelligent systems for mixing multichannel audio," *17th DSP International Conference on Digital Signal Processing*, 2011.
- [12] A. Lerch, *An Introduction to Audio Content Analysis*, Wiley, 2012.
- [13] T. Giannakopoulos and A. Pirkakis, *Introduction to Audio Analysis: A MATLAB Approach*, 1st ed., Academic Press, 2014.