# Bebek Ağlaması Analizinde Multimodal Yaklaşımlar ve Çevresel Bağlam Entegrasyonu: Kapsamlı Literatür Taraması ve Fizibilite Raporu

## Yönetici Özeti

Yenidoğanların temel iletişim mekanizması olan ağlama eyleminin otomatik analizi, son yıllarda dijital sinyal işleme (DSP) ve yapay zeka (AI) teknolojilerindeki gelişmelerle birlikte önemli bir araştırma alanı haline gelmiştir. Ebeveynlere ve sağlık profesyonellerine bebeklerin fizyolojik ve duygusal durumları (açlık, ağrı, yorgunluk, rahatsızlık) hakkında geri bildirim sağlayan bu sistemler, geleneksel olarak yalnızca akustik veriye dayalı "tek modlu" (unimodal) analiz yöntemlerini kullanmaktadır. Bu rapor, "Bebekler İçin 'Neden Ağlıyor?' Tercümanı" projesi kapsamında önerilen **Multimodal (Ses + Sıcaklık/Nem) Entegrasyonu** fikrinin bilimsel temellerini, mevcut pazar boşluğunu ve teknik uygulanabilirliğini değerlendirmek amacıyla hazırlanmıştır.

Mevcut akademik literatür ve ticari ürünler incelendiğinde, bebek ağlaması sınıflandırmasında derin öğrenme modellerinin (CNN, RNN, Transformer) yüksek doğruluk oranlarına ulaştığı görülmektedir. Ancak, bu modellerin büyük çoğunluğu "bağlam körü" (context-blind) olarak çalışmakta; yani bebeğin içinde bulunduğu çevresel koşulları (oda sıcaklığı, nem oranı) analiz sürecine dahil etmemektedir. Ticari ürünlerde (CuboAi, Philips Avent, Owlet) sıcaklık ve nem sensörleri bulunmakla birlikte, bu veriler genellikle yapay zeka algoritmasından bağımsız olarak, yalnızca ebeveyn bilgilendirme ekranında gösterilen statik veriler olarak kullanılmaktadır.

Rapor, önerilen projenin **"Özellik Seviyesinde Füzyon" (Feature-Level Fusion)** mimarisi ile mevcut çözümlerden ayrıştığını tespit etmiştir. Sıcaklık ve nem verilerinin, ağlama sınıflandırma algoritmasına (örneğin bir Konvolüsyonel Sinir Ağı'na) doğrudan bir giriş vektörü olarak dahil edilmesi, özellikle "Rahatsızlık" (Discomfort) sınıfının tespitinde, termal strese bağlı ağlamaları diğer nedenlerden ayırmada özgün ve yenilikçi bir yaklaşım sunmaktadır. Bu çalışma, önerilen multimodal yapının, mevcut "akustik çeviri" ve "ortam izleme" fonksiyonlarını birleştirerek, bebeğin ihtiyaçlarını daha bütüncül bir şekilde yorumlayan "Bağlam Farkındalıklı" (Context-Aware) bir sistem oluşturma potansiyelini detaylandırmaktadır.

## 1. Giriş: Bebek Ağlaması Analizinin (ICA) Evrimi ve Önemi

### 1.1 İletişimsel Bir Sinyal Olarak Ağlama

Bebek ağlaması, yaşamın ilk aylarında bebeğin hayatta kalmasını sağlayan en temel ve ilkel iletişim aracıdır. Bu akustik sinyal, yalnızca rastgele bir gürültü değil; merkezi sinir sistemi, solunum sistemi ve gırtlak kaslarının karmaşık bir koordinasyonu sonucu üretilen, biyolojik belirteçler taşıyan yapılandırılmış bir mesajdır.1 1960'larda Wasz-Höckert ve ekibinin öncü çalışmaları, eğitimli dinleyicilerin (hemşireler ve anneler) bebek ağlamalarını açlık, ağrı, doğum ve zevk gibi kategorilere ayırabildiğini göstermiştir. Ancak bu öznel değerlendirme, dinleyicinin deneyimine, yorgunluğuna ve önyargılarına açık olduğundan, objektif ve tutarlı bir analiz ihtiyacı doğmuştur.

Yenidoğan ağlaması, akustik özellikler açısından yetişkin konuşmasından belirgin farklılıklar gösterir. Bebeklerin vokal yolları daha kısa, gırtlak yapıları daha yüksektedir ve ses üretimi üzerindeki nöromotor kontrolleri henüz tam gelişmemiştir. Bu durum, ağlama sinyallerinin yüksek temel frekans ($F\_0$), ani perde değişimleri ve disharmonik bileşenler içermesine neden olur.2 Bu biyolojik özellikler, otomatik sınıflandırma sistemlerinin tasarımında dikkate alınması gereken en kritik parametrelerdir.

### 1.2 Otomatik Sinyal İşlemeden Derin Öğrenmeye Geçiş

Bebek Ağlaması Sınıflandırması (Infant Cry Classification - ICC) alanı, son yirmi yılda dramatik bir dönüşüm geçirmiştir.

* **Geleneksel Yöntemler:** Erken dönem çalışmalar, ses sinyalinden el yordamıyla çıkarılan özelliklere (Handcrafted features) dayanıyordu. Temel frekans ($F\_0$), enerji konturları ve Mel-Frekans Kepstral Katsayıları (MFCC) gibi öznitelikler, Destek Vektör Makineleri (SVM) veya Gizli Markov Modelleri (HMM) gibi klasik makine öğrenmesi algoritmalarıyla sınıflandırılıyordu.3
* **Derin Öğrenme Çağı:** Günümüzde ise Konvolüsyonel Sinir Ağları (CNN) ve Tekrarlayan Sinir Ağları (RNN) gibi derin öğrenme mimarileri, ham ses verisinden veya spektrogram görüntülerinden özellikleri otomatik olarak öğrenme yeteneğine sahiptir. Bu modeller, "Açlık", "Ağrı", "Yorgunluk" ve "Rahatsızlık" gibi sınıfları %90'ın üzerinde doğrulukla ayırt edebilmektedir.4

### 1.3 "Bağlam Körlüğü" Sorunu ve Multimodal Çözüm İhtiyacı

Mevcut akademik ve ticari çalışmaların büyük çoğunluğu, ağlamayı izole bir ses olayı olarak ele almaktadır. Oysa bir bebeğin ağlama nedeni, çevresel faktörlerden bağımsız değildir. Örneğin, "Rahatsızlık" (Discomfort) kategorisi; ıslak bez, kaşındıran kıyafet veya **uygunsuz oda sıcaklığı** gibi çeşitli fiziksel stres faktörlerinden kaynaklanabilir. Sadece sesi analiz eden bir model, 30°C sıcaklıktaki bir odada ağlayan bebeğin "Termal Stres/Sıcaklama" nedeniyle ağladığını tespit etmekte zorlanabilir; çünkü termal stresin akustik imzası, genel rahatsızlık veya yorgunluk ile karışabilir.

Önerilen proje, bu "Bağlam Körlüğü" (Context Blindness) sorununu çözmek için akustik veriyi (Ses) çevresel veriyle (Sıcaklık/Nem) birleştiren **Multimodal** bir yapı sunmaktadır. Bu yaklaşım, ağlama nedenini tespit ederken dışsal olasılıkları da hesaba katarak, özellikle evde bakım senaryolarında daha güvenilir ve aksiyon alınabilir sonuçlar üretmeyi hedeflemektedir.

## 2. Akademik Literatür Taraması: Mevcut Durum ve Boşluk Analizi

Bu bölümde, bebek ağlaması analizi üzerine yapılan akademik çalışmalar, kullanılan yöntemler ve projenizin önerdiği çevresel entegrasyonun literatürdeki yeri detaylıca incelenmiştir.

### 2.1 Akustik Sınıflandırma Mimarileri

Literatürün ağırlık merkezi, ses sinyallerinin işlenmesi ve sınıflandırılması üzerinedir. Projenizin ses bileşeni için temel teşkil edecek bu çalışmalar, genellikle üç ana aşamadan oluşur: Ön İşleme, Öznitelik Çıkarımı ve Sınıflandırma.

#### 2.1.1 Ön İşleme ve Öznitelik Çıkarımı

Bebek seslerinin yüksek frekanslı yapısı ve gürültüye duyarlılığı nedeniyle ön işleme kritik öneme sahiptir.

* **Vurgulama (Pre-emphasis):** 2 kaynağına göre, bebeklerin ses sistemi özelliklerinden dolayı ağız ve burundan yayılan radyasyonun etkisini dengelemek ve yüksek frekanslı bileşenleri güçlendirmek için ön vurgulama uygulanması şarttır.
* **Çerçeveleme ve Pencereleme:** Ses sinyali durağan olmadığından (non-stationary), 20-30 milisaniyelik kısa çerçevelere bölünür ve spektral sızıntıyı önlemek için Hamming veya Hanning pencereleri ile işlenir.2
* **MFCC (Mel-Frequency Cepstral Coefficients):** Literatürde en baskın öznitelik çıkarma yöntemidir. İnsan kulağının duyma karakteristiğini modelleyen Mel ölçeği kullanılarak elde edilen bu katsayılar, bebek ağlamasının tınısını (timbre) ve fonetik içeriğini temsil etmede son derece başarılıdır. 4 numaralı çalışmada, MFCC tabanlı bir makine öğrenmesi modelinin, Açlık, Rahatsızlık, Karın Ağrısı, Geğirme ve Yorgunluk sınıflarını ayırmada %96.39 gibi çok yüksek bir doğruluğa ulaştığı rapor edilmiştir. Bu, projenizde ses bileşeni için MFCC'nin güçlü bir aday olduğunu doğrulamaktadır.

#### 2.1.2 Derin Öğrenme Modelleri (CNN, RNN, Transformer)

Geleneksel makine öğrenmesinin yerini alan derin öğrenme modelleri, projenizin "Yapay Zeka" katmanı için referans niteliğindedir.

* **Konvolüsyonel Sinir Ağları (CNN):** Görüntü işlemede devrim yaratan CNN'ler, ses spektrogramlarının (sesin zamana bağlı frekans haritası) görüntü olarak işlenmesinde de standart hale gelmiştir. 5 kaynağında, hafifletilmiş (lightweight) bir ResNet mimarisi kullanılarak, kaynak kısıtlı cihazlar (TinyML) için optimize edilmiş bir model geliştirilmiştir. Bu çalışma, Raspberry Pi veya ESP32 gibi gömülü sistemlerde %94 civarında doğrulukla ağlama sınıflandırmasının mümkün olduğunu göstermesi açısından projeniz için kritik bir referanstır.
* **DistilHuBERT ve Transformerlar:** 6 kaynağında bahsedilen DistilHuBERT gibi Transformer tabanlı modeller, büyük veri kümeleri üzerinde eğitilmiş dil modellerinin gücünü ses analizine taşır. Bu modeller, sesin uzun süreli zamansal ilişkilerini yakalamada CNN'lerden daha başarılı olabilir, ancak hesaplama maliyetleri yüksektir. Gömülü sistemlerde (Arduino/ESP32) çalıştırılmaları zor olduğundan, projenizin donanım hedefleri doğrultusunda CNN veya hafifletilmiş RNN modelleri daha uygun görünmektedir.
* **Topluluk Öğrenmesi (Ensemble Learning):** 3 kaynağı, Yapay Sinir Ağları ve Destek Vektör Makineleri topluluklarının (Ensembles) ağrı ve açlık ağlamalarını ayırt etmede tekil modellerden daha başarılı olabileceğini öne sürmektedir.

### 2.2 Multimodal Araştırmalar: Biyomedikal ve Çevresel Ayrımı

"Multimodal" terimi literatürde sıkça geçse de, projenizin odaklandığı "Sıcaklık/Nem" entegrasyonu ile literatürdeki yaygın multimodal çalışmalar arasında önemli bir fark bulunmaktadır.

#### 2.2.1 Biyomedikal Multimodal Füzyon (Yüksek Teknoloji)

Zoundream ve çeşitli üniversite grupları tarafından yürütülen ileri düzey araştırmalar, bebek ağlamasını **biyolojik sinyallerle** birleştirmektedir.

* **Nörofizyolojik Sinyaller:** 7 ve 8 kaynaklarında detaylandırıldığı üzere, ağlama sesleri; EEG (beyin dalgaları), NIRS (beyin oksijenlenmesi), yüz ifadeleri (video analizi) ve vücut hareketleri ile birlikte analiz edilmektedir.
* **Amaç:** Bu çalışmaların temel amacı, otizm spektrum bozukluğu (ASD), işitme kaybı veya merkezi sinir sistemi hastalıkları gibi patolojilerin erken teşhisidir.
* **Projenizle Farkı:** Bu çalışmalar, hastane ortamında ve pahalı sensörlerle (EEG başlıkları, NIRS cihazları) yapılmaktadır. Ev tipi, temassız ve düşük maliyetli bir çözüm sunmayı hedefleyen projenizden kapsam ve donanım gereksinimleri açısından tamamen ayrışmaktadır.

#### 2.2.2 Çevresel Sensör Entegrasyonu ve Literatürdeki Boşluk

Projenizin temelini oluşturan **çevresel sensörlerin (sıcaklık/nem)** kullanımı, literatürde genellikle "füzyon" (birleştirme) değil, "paralel izleme" olarak karşımıza çıkmaktadır.

* **Paralel İzleme Sistemleri:** Birçok akademik çalışma 6, Raspberry Pi veya benzeri kartlar kullanarak "Akıllı Bebek İzleme Sistemi" tasarlamıştır. Bu sistemlerde bir mikrofon ve bir DHT11/DHT22 (Sıcaklık/Nem) sensörü bulunur.
* **Kritik Tespit:** 6 analizinde vurgulandığı üzere, bu çalışmalarda sıcaklık verisi **sınıflandırma algoritmasının bir girdisi olarak kullanılmamaktadır.**
  + *Mevcut Yaklaşım:* Sistem sesi analiz eder -> "Ağlıyor" der. Sistem sıcaklığı ölçer -> "28°C" der. Bu iki bilgi ekranda yan yana gösterilir ama birbirini etkilemez.
  + *Projenizin Farkı:* Sizin önerdiğiniz yapıda (Multimodal Yapı), sıcaklık ve nem verisi, ağlama analizinin **matematiksel bir parçası** olmalıdır. Yani model, "Bu ses örüntüsü normalde 'Açlık' gibi duruyor ama ortam sıcaklığı 32°C olduğu için 'Termal Rahatsızlık' olma ihtimali %80'dir" diyebilmelidir. Literatürde bu tür bir **Özellik Seviyesinde Füzyon (Feature-Level Fusion)** 11 uygulayan bebek izleme çalışması son derece azdır veya deneysel aşamadadır.

### 2.3 Veri Setleri (Datasets) ve Sınırlılıklar

Literatürdeki en büyük zorluk, etiketli ve çevresel veri içeren veri setlerinin eksikliğidir.

| **Veri Seti** | **İçerik** | **Sınıflar** | **Kaynak** | **Proje İçin Değerlendirme** |
| --- | --- | --- | --- | --- |
| **Donate-a-Cry** | ~457 ses dosyası (Kullanıcı yüklemesi) | Açlık, Geğirme, Ağrı, Yorgunluk, Rahatsızlık | 12 | GitHub'da açık kaynaklıdır. En popüler veri setidir ancak etiketlerin doğruluğu şüphelidir ve sınıflar dengesizdir (Çoğunlukla "Açlık"). Çevresel veri içermez. |
| **Baby Chillanto** | ~2,268 ses klibi (1'er saniye) | Asfiksi, Sağır, Açlık, Normal, Ağrı | 2 | Meksika kaynaklıdır. Patoloji tespiti (sağırlık vb.) odaklıdır. Günlük ihtiyaç analizi için süreleri çok kısadır. |
| **Dunstan Baby Language** | Video/Ses kesitleri | Neh (Aç), Owh (Yorgun), Eh (Gaz), Eairh (Karın), Heh (Rahatsız) | 15 | Priscilla Dunstan'ın teorisine dayanır. Bilimsel geçerliliği tartışmalıdır ancak sınıf çeşitliliği açısından değerlidir. |
| **ESC-50** | Çevresel Sesler | Sınıf #20: Ağlayan Bebek | 14 | Yüksek kaliteli ama sadece tek bir sınıf ("Ağlama") içerir. "Neden ağlıyor?" sorusuna cevap veremez, sadece "Ağlıyor mu?" sorusunu yanıtlar. |

**Özgünlük Fırsatı:** Literatürde "Ses + Sıcaklık" etiketli bir veri seti bulunmamaktadır. Bu durum, projenizin veri toplama veya sentetik veri üretme (mevcut ses verilerine sıcaklık etiketi ekleyerek model eğitme) aşamasında özgün bir katkı sunabileceğini göstermektedir.

## 3. Ticari Ürünler ve Pazar Analizi

Mevcut pazar, donanım tabanlı "Akıllı Monitörler" ve yazılım tabanlı "Çeviri Uygulamaları" olarak ikiye ayrılmaktadır. Projenizin konumu bu iki grubun kesişim noktasındadır.

### 3.1 Zoundream (Teknoloji Lideri ve Araştırma Odaklı)

İsviçre merkezli Zoundream 18, bu alandaki en gelişmiş teknoloji sağlayıcısıdır.

* **Teknoloji:** **AMSI (Acoustic Multi-Stage Interpreter)** adını verdikleri derin öğrenme motorunu kullanırlar. Bebek ağlamasını bir sağlık biyobelirteci olarak ele alırlar.
* **İş Modeli:** Doğrudan tüketiciye ürün satmak yerine, teknolojilerini **Maxi-Cosi** gibi donanım üreticilerine lisanslarlar.
* **Multimodal Yaklaşım:** Zoundream, EEG ve NIRS gibi biyomedikal verilerle multimodal çalışmaları aktif olarak yürütmektedir.18 Ancak ticari ürünlerinde (örneğin Maxi-Cosi See Pro 360 20), çevresel sensörler (sıcaklık/nem) bulunsa da, bu sensörlerin ağlama "çeviri" algoritmasına dahil edildiğine dair bir kanıt yoktur. "CryAssist" teknolojisi sesi analiz ederken, cihazdaki sensörler ebeveynlere oda sıcaklığını göstermek için kullanılmaktadır.

### 3.2 Cappella (Mobil Uygulama Girişimi)

ABD merkezli Cappella 21, abonelik tabanlı bir mobil uygulama sunmaktadır.

* **İddia:** %95 doğrulukla ağlama nedenini (Açlık, Rahatsızlık, Yorgunluk vb.) tespit ettiğini belirtmektedir.
* **Sıcaklık Entegrasyonu:** CES 2024 tanıtımlarında ve basın bültenlerinde, Cappella yetkilileri *"Sıcaklık üzerinde hala çalışıyoruz -- bebeğinizin çok sıcak veya soğuk olduğu için ağlayıp ağlamadığını analiz etmek için bir başka veri noktası olarak"* ifadesini kullanmıştır.21
* **Fırsat Tespiti:** Bu ifade, projenizin odaklandığı noktanın (sıcaklığın analiz için bir veri noktası olması) **henüz ticari olarak tam çözülmemiş, aktif bir Ar-Ge konusu olduğunu** kanıtlamaktadır. Büyük bir girişimin "üzerinde çalışıyoruz" dediği bir özelliği projenizde hayata geçirmek, çalışmanızı son derece güncel ve rekabetçi kılar.

### 3.3 Geleneksel Akıllı Monitörler (CuboAi, Owlet, Philips)

* CuboAi 23: Bilgisayarlı görü (Computer Vision) üzerine yoğunlaşmıştır. Yüzün örtülmesi, dönme hareketleri gibi güvenlik unsurlarını takip eder. Sıcaklık sensörü sadece uyarı amaçlıdır.
* Owlet 24: "Smart Sock" ile nabız ve oksijen satürasyonu takibi yapar.
* Philips Avent 25: Zoundream teknolojisini kullanarak ağlama çevirisi yapar ancak donanım entegrasyonu sınırlıdır.
* **Eksiklik:** Bu cihazlar genellikle "Uyarıcı" sistemlerdir. "Bebek ağlıyor mu?" ve "Oda sıcak mı?" sorularını ayrı ayrı yanıtlarlar. Sizin projeniz ise "Bebek oda sıcak olduğu için mi ağlıyor?" sorusunu yanıtlayarak nedensellik bağı kurmaktadır.

## 4. Açık Kaynak Projeler ve Teknik Uygulanabilirlik (GitHub/TinyML)

Projenizin akademik veya prototip aşaması için açık kaynak ekosistemi zengin kaynaklar sunmaktadır. Özellikle **TinyML** (Mikrodenetleyicilerde Makine Öğrenmesi) alanı, projenizin donanım hedefleriyle birebir örtüşmektedir.

### 4.1 Edge Impulse ve TinyML Projeleri

GitHub ve Edge Impulse platformunda yapılan incelemelerde 26, benzer donanımlar kullanan birçok proje tespit edilmiştir.

* **Mevcut Projeler:** Genellikle **Arduino Nano 33 BLE Sense** kartı kullanılmaktadır. Bu kartın seçilme nedeni, üzerinde hem dijital mikrofonun (MP34DT05) hem de sıcaklık/nem sensörünün (HTS221) dahili olarak bulunmasıdır.
* **Sınırlılık:** Mevcut açık kaynak projelerin %90'ı **"Ağlama Tespiti" (Cry Detection)** üzerine odaklıdır; yani sadece "Gürültü mü, Bebek Ağlaması mı?" ayrımı yaparlar (İkili Sınıflandırma). "Neden Ağlıyor?" (Açlık/Ağrı/Rahatsızlık) ayrımı yapan projeler 29 ise veri yetersizliğinden dolayı düşük doğruluk oranlarında kalmaktadır ve sadece sesi kullanmaktadır.
* **Sensör Füzyonu Desteği:** Edge Impulse platformu, yakın zamanda **"Sensor Fusion"** (Sensör Füzyonu) özelliğini tanıtmıştır.30 Bu özellik, farklı sensörlerden gelen verilerin (örneğin ivmeölçer + ses) tek bir sinir ağına beslenmesine olanak tanır. Projeniz, bu yeni özelliği "Ses + Sıcaklık" ikilisine uyarlayarak açık kaynak topluluğuna özgün bir katkı sağlayabilir.

### 4.2 Donanım ve Yazılım Altyapısı

Projenizin teknik altyapısı için literatürdeki en uygun bileşenler şunlardır:

* **İşlemci:**
  + **Raspberry Pi 4/5:** Yüksek işlem gücü, Python desteği ve kolay veri toplama için uygundur.5 Ancak enerji tüketimi yüksektir.
  + **ESP32 / Arduino Nano 33 BLE Sense:** Düşük güç tüketimi ve TinyML desteği ile "Uçta İşleme" (Edge Computing) için idealdir. Gizlilik odaklı (sesin buluta gitmediği) bir ürün için bu kartlar tercih edilmelidir.32
* **Sensörler:**
  + **BME680/BME280:** DHT11/22 serisine göre çok daha hassas ve hızlı tepki süresine sahiptir. "Bağlam Farkındalıklı" bir AI için ani sıcaklık değişimlerini yakalamak önemlidir.
* **Yazılım Kütüphaneleri:** TensorFlow Lite for Microcontrollers (TFLite Micro) ve Edge Impulse SDK.

## 5. Projenizin Farkı ve Özgün Tarafları (Differentiation)

Kullanıcı sorgusunda özellikle talep edilen "bizim fikrimizin farkı" konusu, literatür ve pazar analizi ışığında aşağıdaki başlıklarda netleştirilmiştir:

### 5.1 Entegrasyon Seviyesi: Füzyon vs. Paralellik

Mevcut ürünler ve projeler **"Paralel İzleme"** (Parallel Monitoring) yaparken, sizin projeniz **"Algoritmik Füzyon"** (Algorithmic Fusion) önermektedir.

* *Mevcut:* "Ağlama Sesi Algılandı" uyarısı gönder. Ayrıca "Oda Sıcaklığı 26°C" bilgisini göster. Kararı ebeveyne bırak.
* *Özgün Taraf (Sizin Projeniz):* Sıcaklık verisini bir katsayı veya özellik vektörü olarak yapay zeka modeline besle. Model, oda sıcaklığının yüksek olduğu durumlarda akustik özellikler belirsiz olsa bile "Rahatsızlık/Sıcaklama" sınıfının olasılık puanını matematiksel olarak artırır. Bu, ebeveynin bilişsel yükünü azaltan bir "Karar Destek Sistemi"dir.

### 5.2 "Termal Rahatsızlık" Alt Sınıfı

Literatürdeki "Rahatsızlık" (Discomfort) sınıfı çok geniştir (ıslak bez, gaz, sıcak, soğuk, kaşıntı). Sizin projeniz, çevresel sensörleri kullanarak bu sınıfı daha spesifik hale getirebilir.

* *Özgün Taraf:* Projeniz, genel bir "Rahatsızlık" uyarısı yerine, ortam verisiyle desteklenmiş **"Ortam Kaynaklı Rahatsızlık"** (Environmental Discomfort) teşhisi koyabilir. Bu, ebeveynin sorunu çözme süresini (pencereyi açmak veya kombiyi kısmak gibi) hızlandırır.

### 5.3 Gizlilik ve Uçta İşleme (Edge AI)

Ticari rakiplerin çoğu (Cappella, Philips/Zoundream) karmaşık analizleri yapmak için bulut sunucularını kullanabilir veya yüksek işlem gücü gerektirebilir.

* *Özgün Taraf:* Önerilen TinyML yaklaşımı ve sensör füzyonu, düşük işlem gücüyle (Arduino/ESP32) yüksek çıkarım yeteneği sağlar. Ses verisi cihazdan hiç çıkmaz, analiz cihaz üzerinde (on-device) yapılır. Bu, "Veri Gizliliği" hassasiyeti olan ebeveynler için büyük bir fark yaratır.

### 5.4 Özet Karşılaştırma Tablosu

| **Özellik** | **Akademik Çalışmalar** | **Ticari Ürünler (Zoundream/Cappella)** | **Sizin Projeniz (Multimodal)** |
| --- | --- | --- | --- |
| **Girdi Verisi** | Sadece Ses (Genellikle) | Ses + (Bazen Video/Nabız) | **Ses + Sıcaklık + Nem** |
| **Sıcaklık Rolü** | Yok veya Sadece Ekran Gösterimi | Ekran Gösterimi / Uyarı | **Sınıflandırma Algoritması Girdisi** |
| **Yöntem** | CNN, RNN (Unimodal) | Derin Öğrenme (Bulut/Mobil) | **Sensör Füzyonu (Edge/TinyML)** |
| **Hedef Sınıf** | Genel İhtiyaçlar (Açlık, Ağrı) | Genel İhtiyaçlar + Patoloji | **Bağlam Destekli İhtiyaç Analizi** |

## 6. Teknik Uygulama Yol Haritası ve Öneriler

Projenin hayata geçirilmesi ve literatürdeki boşluğu doldurması için aşağıdaki teknik adımlar önerilmektedir:

### 6.1 Veri Seti Mühendisliği (Sentetik Füzyon)

Açık kaynaklı bir "Ses + Sıcaklık" veri seti bulunmadığından, eğitim verisi şu şekilde oluşturulmalıdır:

1. **Ses Kaynağı:** *Donate-a-Cry* veri setindeki "Rahatsızlık" (Discomfort), "Açlık" (Hunger) ve "Ağrı" (Pain) sınıfları temizlenerek kullanılır.
2. **Etiket Zenginleştirme (Sentetik):**
   * "Açlık" ve "Ağrı" ses dosyalarına rastgele "Normal Oda Koşulları" (20-24°C) etiketleri atanır.
   * "Rahatsızlık" dosyalarının bir kısmı çoğaltılır ve bunlara "Yüksek Sıcaklık" (>26°C) veya "Yüksek Nem" (>%70) etiketleri atanır.
3. **Amaç:** Modelin, "Rahatsızlık" sesi ile "Yüksek Sıcaklık" verisi arasındaki korelasyonu (ilişkiyi) öğrenmesini sağlamaktır.

### 6.2 Model Mimarisi: Özellik Seviyesinde Füzyon (Feature-Level Fusion)

11 kaynağında belirtilen teoriye uygun olarak:

1. **Ses Kolu:** Ses sinyali MFCC'ye dönüştürülür ve 1D-CNN (veya 2D-CNN) katmanlarından geçirilerek bir özellik vektörü (embedding) elde edilir.
2. **Sensör Kolu:** Sıcaklık ve Nem verileri normalleştirilir (0-1 arası) ve basit bir Tam Bağlı (Dense) katmandan geçirilir.
3. **Birleştirme (Concatenation):** Ses ve sensörden gelen vektörler birleştirilir.
4. **Sınıflandırıcı:** Birleşik vektör, son bir Softmax katmanına girer ve sonuç üretilir.

Bu mimari, Edge Impulse platformunda "Keras Expert Mode" kullanılarak veya Python (TensorFlow/Keras) ile kodlanarak Arduino Nano 33 BLE Sense üzerinde çalıştırılabilir.

### 6.3 Sonuç

Önerilen "Multimodal Bebek Ağlaması Tercümanı", literatürdeki **"Bağlam Körlüğü"** sorununa doğrudan bir çözüm sunmaktadır. Ticari rakiplerin (Cappella) henüz Ar-Ge aşamasında olduğu "sıcaklık entegrasyonu" özelliğini, donanım seviyesinde (IoT) ve özellik füzyonu mimarisiyle hayata geçirmek, projenizi hem akademik hem de ticari açıdan özgün ve değerli kılmaktadır.

### Referanslar (Metin İçinde Kullanılan Kaynak Kodları)

* **Genel İnceleme ve Yöntemler:**.1
* **Veri Setleri:** 12 (Donate-a-Cry)2 (Baby Chillanto)15 (Dunstan)17 (ESC-50).
* **Ticari Analiz:** 18 (Zoundream)21 (Cappella)23 (CuboAi).
* **Multimodal ve Füzyon:** 7 (Biyomedikal)11 (Füzyon Teorisi)30 (Edge Impulse).
* **Donanım ve TinyML:**.5

#### Alıntılanan çalışmalar

1. Advances in Infant Cry Paralinguistic Classification—Methods, Implementation, and Applications: Systematic Review - NIH, erişim tarihi Aralık 26, 2025, <https://pmc.ncbi.nlm.nih.gov/articles/PMC12076029/>
2. Classification of Infant Crying Sounds Using SE-ResNet-Transformer - MDPI, erişim tarihi Aralık 26, 2025, <https://www.mdpi.com/1424-8220/24/20/6575>
3. Identifying Pain and Hunger in Infant Cry with Classifiers Ensembles - ResearchGate, erişim tarihi Aralık 26, 2025, <https://www.researchgate.net/publication/4242623_Identifying_Pain_and_Hunger_in_Infant_Cry_with_Classifiers_Ensembles>
4. Machine learning-based infant crying interpretation - Frontiers, erişim tarihi Aralık 26, 2025, <https://www.frontiersin.org/journals/artificial-intelligence/articles/10.3389/frai.2024.1337356/full>
5. TinyML Prototype of Infant Cry Classification System - IEEE Xplore, erişim tarihi Aralık 26, 2025, <https://ieeexplore.ieee.org/iel8/6287639/10820123/11132295.pdf>
6. Development of a Baby Cry Identification System Using a Raspberry Pi-Based Embedded System and Machine Learning - MDPI, erişim tarihi Aralık 26, 2025, <https://www.mdpi.com/2227-7080/13/4/130>
7. The significance of an infant's cry: a narrative review of physiological, pathological, and analytical perspectives - PubMed Central, erişim tarihi Aralık 26, 2025, <https://pmc.ncbi.nlm.nih.gov/articles/PMC12098436/>
8. Publications - Zoundream, erişim tarihi Aralık 26, 2025, <https://zoundream.com/publications/>
9. Design of a hazard prediction system with intelligent multimodal fusion based on artificial intelligence & internet of things technology: taking a crib as an example - PubMed Central, erişim tarihi Aralık 26, 2025, <https://pmc.ncbi.nlm.nih.gov/articles/PMC11623082/>
10. A novel smart baby cradle system utilizing IoT sensors and machine learning for optimized parental care - PubMed Central, erişim tarihi Aralık 26, 2025, <https://pmc.ncbi.nlm.nih.gov/articles/PMC12125208/>
11. Contactless Vital Sign Monitoring: A Review Towards Multi-Modal Multi-Task Approaches, erişim tarihi Aralık 26, 2025, <https://pmc.ncbi.nlm.nih.gov/articles/PMC12349365/>
12. Donate a cry corpus Dataset Distribution. | Download Scientific Diagram - ResearchGate, erişim tarihi Aralık 26, 2025, <https://www.researchgate.net/figure/Donate-a-cry-corpus-Dataset-Distribution_tbl1_384064081>
13. Machine learning-based infant crying interpretation - PMC - NIH, erişim tarihi Aralık 26, 2025, <https://pmc.ncbi.nlm.nih.gov/articles/PMC10882089/>
14. ICSD: An Open-source Dataset for Infant Cry and Snoring Detection - arXiv, erişim tarihi Aralık 26, 2025, <https://arxiv.org/html/2408.10561v1>
15. Infant Cry Analysis: A Survey of Datasets, Features, and Machine Learning Techniques - LJMU Research Online, erişim tarihi Aralık 26, 2025, <https://researchonline.ljmu.ac.uk/id/eprint/27659/1/Infant%20Cry%20Analysis%20A%20Survey%20of%20Datasets%20Features%20and%20Machine%20Learning%20Techniques.pdf>
16. (PDF) Baby Cry Sound Detection: A Comparison of Mel Spectrogram Image on Convolutional Neural Network Models - ResearchGate, erişim tarihi Aralık 26, 2025, <https://www.researchgate.net/publication/384064081_Baby_Cry_Sound_Detection_A_Comparison_of_Mel_Spectrogram_Image_on_Convolutional_Neural_Network_Models>
17. karolpiczak/ESC-50: ESC-50: Dataset for Environmental ... - GitHub, erişim tarihi Aralık 26, 2025, <https://github.com/karolpiczak/ESC-50>
18. Zoundream: Home, erişim tarihi Aralık 26, 2025, <https://zoundream.com/>
19. Technology - Zoundream, erişim tarihi Aralık 26, 2025, <https://zoundream.com/technology/>
20. See Pro 360° Baby Monitor - Maxi-Cosi, erişim tarihi Aralık 26, 2025, <https://maxicosi.com/products/see-pro-360-baby-monitor-mo183>
21. This App Says It Can Translate Your Baby's Cries Using AI - BusinessGhana, erişim tarihi Aralık 26, 2025, <https://www.businessghana.com/site/news/Technology/299989/This-App-Says-It-Can-Translate-Your-Baby's-Cries-Using-AI->
22. AI tool claims to translate what your baby's cries mean | HerFamily IE, erişim tarihi Aralık 26, 2025, <https://www.herfamily.ie/parenting/ai-tool-claims-to-translate-what-your-babys-cries-mean-444517>
23. CuboAi Smart Baby Monitor With Breathing Detection - 2.5K HD Camera & Sleep Safety Alerts, erişim tarihi Aralık 26, 2025, <https://www.talbrossealing.com/Baby-Monitor-With-Breathing-Detection-2-5K-HD-Camera-Sleep-634232>
24. Owlet Care - App Store, erişim tarihi Aralık 26, 2025, <https://apps.apple.com/bg/app/owlet-care/id1343309759>
25. How does the Cry Translation feature work? | Avent - Philips, erişim tarihi Aralık 26, 2025, <https://www.usa.philips.com/c-f/XC000021720/how-does-the-cry-translation-feature-work>
26. Smart Baby Swing - Arduino Portenta H7 - Edge Impulse Documentation, erişim tarihi Aralık 26, 2025, <https://docs.edgeimpulse.com/projects/expert-network/smart-baby-swing-arduino-portenta-h7>
27. BABL: A Baby Monitor Powered by tinyML and Edge Impulse! - Hackster.io, erişim tarihi Aralık 26, 2025, <https://www.hackster.io/ishotjr/babl-a-baby-monitor-powered-by-tinyml-and-edge-impulse-f5045f>
28. mandymadongyi/Baby-Cry-Detector-CASA0018: CASA0018 Deep Learning for Sensor Networks Individual Project - GitHub, erişim tarihi Aralık 26, 2025, <https://github.com/mandymadongyi/Baby-Cry-Detector-CASA0018>
29. Baby Cry Audio Data Strategy Guidance - Edge Impulse Forum, erişim tarihi Aralık 26, 2025, <https://forum.edgeimpulse.com/t/baby-cry-audio-data-strategy-guidance/10039>
30. Classify multiple 2D input features - Edge Impulse Documentation, erişim tarihi Aralık 26, 2025, <https://docs.edgeimpulse.com/tutorials/topics/machine-learning/classify-multiple-2d-features>
31. Environmental sensor fusion - Edge Impulse Documentation, erişim tarihi Aralık 26, 2025, <https://docs.edgeimpulse.com/tutorials/end-to-end/environmental-sensor-fusion>
32. (PDF) LimitAccess: on-device TinyML based robust speech recognition and age classification - ResearchGate, erişim tarihi Aralık 26, 2025, <https://www.researchgate.net/publication/368759585_LimitAccess_on-device_TinyML_based_robust_speech_recognition_and_age_classification>
33. Baby Language – Apps on Google Play, erişim tarihi Aralık 26, 2025, <https://play.google.com/store/apps/details?id=com.enbyin.babylanguage&hl=en_IN>
34. Sensor fusion algorithms | imec, erişim tarihi Aralık 26, 2025, <https://www.imec-int.com/en/79GHz-140GHz-radar-systems/sensor-data-fusion/sensor-fusion-algorithms>