

Açıklanabilir Yapay Zeka Katmanlı V-Model Otonom Sürüş Yazılım Mimarisi

Explainable Artificial Intelligence Layered V-Model Autonomous Driving Software Architecture

*Musa Nurullah Yazar¹, Berk Ađın¹, Pelin Öksüz¹, Candaş Ünal¹, Cemal Ege Özdemir¹,
Melik Buğra Özçelik¹, Uygur Kaya¹*

¹AVL Türkiye Araştırma ve Mühendislik

{musa.yazar, berk.agin, pelin.oksuz, candas.unal, ege.ozdemir, melik.ozcelik,
uygar.kaya}@avl.com

Özetçe

Son yıllarda yapay zeka ve özellikle derin öğrenme yöntemlerindeki gelişmelerden faydalanarak araçların bulunduğu dinamik çevreyi insan düzeyinde algılayıp yorumlamasını ve kendi kararlarını verebilmesini sağlayan otonom sürüş yazılım sistemlerinin geliştirilmesine yönelik çalışmalar yoğun bir şekilde devam etmektedir. Ayrıca yakın gelecekte yaygınlaşması beklenen yüksek güvenilirlikli ve yüksek hızlı ağ sistemleri ile araçların birbirleriyle ve çevreyle haberleşme yetkinliğine sahip olacağı, böylece araçların daha güvenli ve ekonomik bir sürüş sağlayan kooperatif sistemlere dönüşeceği öngörülmektedir. Otonom sürüş yazılım sistemlerinde kullanılan yapay zeka tabanlı sistemlerin açıklanabilir, şeffaf, güvenilir ve standartlara uyumlu olması beklenmektedir. Bu çalışmada yazılım dünyasında yoğunlukla kullanılan V-model mimariye şekilsel ve işlevsel benzerlikleri olan beş katmanlı bir otonom sürüş yazılım mimarisi ele alınmış ve yazılım mimarisinin otomotiv sektöründe yukarıda ifade edilen isterleri karşılayabilmesi hedeflenmiştir. Ayrıca son yıllarda geliştirilen açıklanabilir yapay zeka (AYZ) ve derin öğrenme yöntemlerinin otonom sürüş yazılım mimarisinin ilgili katmanındaki yeri ve diğer katmanlarla olan ilişkisi detaylı bir şekilde sunulmuştur.

Abstract

In recent years, there have been intensive efforts to develop autonomous driving software systems that enable vehicles to perceive and interpret the dynamic environment at the human level and make their own decisions by utilizing the developments in artificial intelligence and especially deep learning methods. In addition, with the high-security and high-speed network systems that are expected to become widespread in the near future, vehicles will have the ability to communicate with each other and the environment, thus transforming vehicles into cooperative systems that provide safer and more economical driving. Artificial intelligence-based systems used in autonomous driving software systems are expected to be explainable, transparent, reliable and compliant with standards. In this study, a

five-layer autonomous driving software architecture, which has morphological and functional similarities to the V-model software architecture, is discussed and it is aimed that the software architecture can meet the above-mentioned requirements in the automotive industry. In addition, the place of explainable artificial intelligence (xAI) and deep learning methods developed in recent years in the relevant layer of the autonomous driving software architecture and their relationship with other layers are presented in detail.

1. Giriş

Otomotiv endüstrisinin önemli bir dönüşümün içinde olduğu, 2030'lu yıllar ve sonrasında yazılımın araç ürün farklılaşmasında ve maliyetinde en önemli faktörlerden biri haline geleceği öngörülmektedir [1]. Günümüz araç teknolojisinde içten yanmalı motorlardan elektrikli ve hibrit araçlara, manuel sürüşten, ileri sürüş destek sistemleri ve farklı seviyelerde otonom sürüş özelliği sağlayan araç modellerine kadar çok sayıda varyasyon bulunmakta ve bu da yeni güvenlik, konfor, performans ve eğlence özelliklerine yönelik geliştirilmesi gereken elektronik kontrol ve yazılım bileşenlerin ağırlığına oldukça artırmaktadır. Özellikle farklı seviyelerde otonom sürüş sağlayan araçlar göz önüne alındığında elektronik kontrol ve yazılımın bileşenlerinin ağırlığının büyük oranda artmaya devam edeceği açıkça görülmektedir. Otonom araç, çoklu algılayıcı füzyonu ve bilgisayar görüşü, olasılık ve karar teorisi, yapay zeka, kontrol teorisi, robotik ve diğer araştırma alanlarının kavramlarını kullanan çeşitli yazılım bileşenleri ve algoritmaları içeren kompleks bir yapıdan oluşmaktadır. Otonom sürüş yazılım mimarisi yukarıda ifade edilen kompleks yapıya ait sistem bileşenlerinin yüksek performansta ve uyumla çalışabilmesini sağlayarak otonom araçların güvenilirliğini, performansını ve kullanıcı deneyimini iyileştirmeyi hedeflemektedir. Günümüzde farklı araştırma grupları tarafından birçok mimari tasarım yaklaşımı önerilmiştir [2, 3]. Genel olarak topolojik ve hiyerarşik sınıflandırmaya dayalı farklı sayıda katmandan oluşan yapılar ön plana çıkmaktadır. Bu katmanlar; algılama (perception), karar verme (decision maker), planlama (planner), araç kontrolü, sistem yönetimi ve insan makine arayüzü (Human-Machine Interface, HMI) şeklinde ele alınmaktadır. Son yıllarda yapay zeka ve derin öğrenme yöntemlerine dayalı algoritmaların algılama ve

karar verme modüllerinde kullanımının artması ile daha başarılı performans ortaya koyabilen sistemler geliştirilebilse de çok katmanlı yapay sinir ağlarına dayalı bu algoritmaların kara-kutu olarak adlandırılan davranışı sistem güvenilirliğini olumsuz yönde etkilemektedir.

Bu çalışmada, otomotiv endüstrisinin yazılım tasarım ilkelerini ve izlenebilirlik gereksinimlerini karşılayan modüler bir mimari yaklaşımı [4] ele alınmış ve sistemin açıklanabilirliğini, güvenilirliğini ve standartlara uyumlu olmasını artıracak açıklanabilir yapay zeka katmanlı V-model otonom sürüş yazılım mimarisi sunulmuştur. Birinci bölümde yazılımın araç ürün farklılaşmasındaki ağırlığı ve otonom sürüş yazılım mimarisinin gereksinimleri üzerine durulmuştur. İkinci bölümde açıklanabilir yapay zeka yöntemlerinin algılama, çevre modeli (world model) ve yorumlayıcı (interpreter), planlama gibi modüllerde kullanımına yönelik literatürde yer alan güncel çalışmalar verilmiştir. Üçüncü bölümde önerilen sistem mimarisi detaylı bir şekilde sunulmuş ve dördüncü bölümde ise sonuç ve öneriler paylaşılmıştır.

2. Literatür Özeti

Açıklanabilir yapay zeka, herhangi bir sistemde kullanılan yapay zekanın davranışlarını insanlara anlatmak için uygulanan bir yöntemdir [5, 6]. Otonom araçlarda kullanılan açıklanabilir yapay zeka ise otonom sürüş amacıyla kullanılan birçok yapay zeka modüllerini geliştiricilere veya son kullanıcı olan sürücülere açıklar ve sistem ile insan arasındaki iletişimi artırarak hem güvenilirliği yükseltir, hem sistemi daha anlaşılır kılar, hem de insanların yönetip geliştirebileceği bir ortam sağlar. Açıklamalar, eylemlerin nedenini, işleyişini ve olası sonuçlarını içerir ve kişilere bu açıklamalar görsel, sözel veya metin olarak sunulabilir. Bu açıklamalar sayesinde aracın kararları ve tepkileri anlaşılabilir hale gelir. Açıklanabilir yapay zekanın tüm bu özelliklerinin yanında, bu metodun eklendiği çalışmalarda eylem tahminlerinin doğruluğunun arttığı gözlenmiştir [7, 8].

Algılama, aracın doğru ve güvenli bir şekilde hareket etmek için otonom aracın çevresindeki nesnelere ve yolları anlamasını sağlayan önemli bir adımdır. Algılama algoritmaları, radar, LIDAR, kameralar ve ultrasonik sensörler gibi çeşitli algılama sensörlerini kullanarak çevreden gelen verileri analiz edip kullanarak nesnelere konumunu, hızını ve diğer özelliklerini belirler. Literatürde, algılama adımında kullanılan yapay zeka modülünün çalışırken kamera görüntülerini kullanarak yoldaki cisimlerin önemini araç sürücülerine veya geliştiricilere gösterebilen çalışmalar mevcuttur. Bunun için görüntü dikkat mekanizmaları ile gerçek zamanlı görüntülerde vurgulanması gereken bölgeler öne çıkarılarak görüntü temizlenir. Bu yaklaşımın otonom araçların performansını ve açıklanabilirliğini olumlu yönde etkilediği gözlemlenmiştir [9, 10, 11].

Çevre modeli, otonom aracın sensörler yardımıyla algıladığı bilgileri anlamlandıran ve yol koşulları hakkında tahminlerde bulunan bir adımdır [12]. Bu model, gelen verileri işler ve aracın çevredeki nesnelere hareketlerini, yol şartlarını ve diğer önemli faktörleri tahmin etmesine yardımcı olur, aracın güncel durumu ve çevrenin dinamikleri hakkında bir anlayış oluşturur. Çevre modeline ek olarak yapay zeka ile çalışan birçok yapı, çalışmasının sonucunda ortaya çıkan bilgiler eyleme geçmeden önce bir anlama sahip olması gerekir. Bunu gerçekleştiren sis-

temler de yorumlayıcı olarak adlandırılır. Yorumlayıcılar gelecek öngörüsünde bulunur ve bunu geçmişten gelen verilerle sağlar. Daha önce yapılmış olan çalışmalarda da ilgili ortamın modeli oluşturulup yorumlayıcıların kullanıldığı algoritmalarda açıklanabilir yapay zeka teknolojilerinden yararlanılmıştır [13, 14, 15].

Planlayıcı modülü, otonom araçlar için kritik bir sistemdir. Şehir içi yollar veya diğer araçların bulunduğu otoyollar gibi karmaşık ortamlarda zor kararlar vermeli ve karmaşık manevralar gerçekleştirmelidir. Son yıllarda, karar verme mekanizması için pekiştirmeli öğrenme algoritmaları üzerinde çalışmalar yapılmaktadır. Bu algoritmalarla eğitilmiş bir etmen (agent), çevre modelini algılayabilir, anlayabilir, yorumlayıcıdan gelen bilgileri kullanabilir ve alınan kararları alt seviye kontrolcüsüne iletebilir. Ancak, karar alma sürecinde alınan kararlar, bu süreci değerlendiren bir kara-kutu modelinden çıkar. Son yıllarda, araştırmacılar etmenin aldığı kararların nedenini açıklayabilme üzerine odaklanmışlardır. Bu, karar verme sürecinin şeffaflığını artırarak, etmenin kararlarının anlaşılabilirliğini ve güvenilirliğini sağlamayı amaçlamaktadır.

Açıklayıcı yapay zeka teknikleri ve yorumlanabilir pekiştirmeli öğrenme yöntemleri güncel çalışmalarda sıkça kullanılmaya başlamıştır. Bu teknikler, etmenin kararlarını insanlar tarafından anlaşılabilir şekilde açıklayabilmesini sağlar. Bu sayede, etmen tarafından alınan kararlar daha iyi değerlendirilebilir ve gerektiğinde müdahale edilebilir. Bu gelişmeler, otonom araçların güvenliğini artırmak, karar alma süreçlerini optimize etmek ve insanlarla etkileşimlerini daha etkili hale getirmek için önemli adımlardır. Gelecekte, açıklanabilir yapay zeka ve pekiştirmeli öğrenme yöntemlerinin daha da ilerlemesi beklenmektedir [16, 17, 18]. [16] çalışmasında, etmen hız sınırlarının olduğu bir ortamda, aldığı kararları algıladığı hız sınırlarını açıklayarak alır. [17] çalışmasında, problem üzerindeki açıklamaların kullanıcıya güven verdiği ve performansı artırdığı ortaya konulmuştur. [18] çalışmasında, trafiğe yeni katılan araçların kararlarını etkileyecek çevre araçlar için yörünge seçim problemi üzerine yoğunlaşmış ve bu araçların yörüngeleri için açıklanabilir güvenilirlik değerleri ortaya konulmuştur.

Gelişmiş teknolojilerin ortaya çıkması ve otonom araçların yükselişiyle birlikte kablosuz iletişim teknolojisi (örneğin, araçtan nesnelere haberleşme, V2X), araçlar ve çevre arasında kesintisiz iletişim ve işbirliği sağlayarak, otonom sürüş esnasında güvenliği ve verimliliği artırmaktadır. Araçların hız, konum, yörünge vb. gibi gerçek zamanlı bilgi alışverişinde bulunmasını sağlayan V2X bağlantısı, koordineli eylemlere ve proaktif karar almaya olanak tanır. Llatser ve diğerlerinin [19] çalıştığı kolektif algı (collective perception) yönteminde yakındaki araçlar arasında sensör verilerinin alışverişi ile beraber otonom araç, diğer araçlar tarafından algılanan nesnelere kendi çevre modellerine dahil etmesiyle farkındalığını artırarak kör noktaları ortadan kaldırır ve güvenilirliğini artırır. Ayrıca, aynı çalışmanın işbirlikçi manevra koordinasyonu (cooperative maneuver coordination) yöntemiyle tüm trafik katılımcılarına fayda sağlayan ortak bir manevra bulmak için araçların sürüş planlarını veya niyetlerini paylaşmasıyla aralarındaki yanlış anlaşılmaları önler ve böylece olası çarpışma risklerinden kaçınırken kaynak kullanımını optimize ederek yoldaki araba sayısını azaltır ve trafik sıkışıklığını hafifletilmesine yardımcı olarak daha verimli bir ulaşım ağına olanak sağlar. Sonuç olarak, bağlantı modülleri (V2X bağlantısı), otonom sürüşün önemli bir parçasıdır çünkü

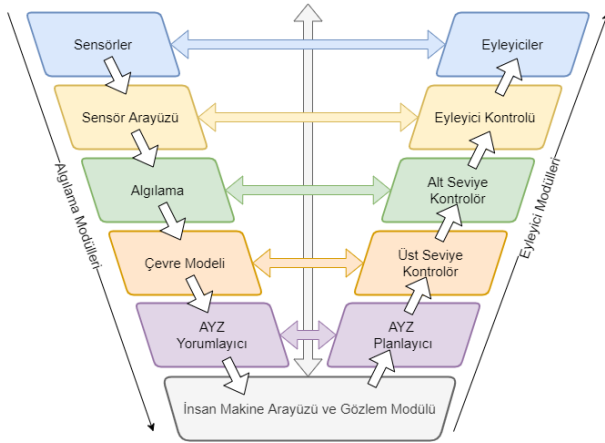
yol güvenliğini artırır, trafik yönetimini optimize eder, ortak hareketliliği destekleyerek sürücüler arasındaki sosyal etkileşimi artırır.

3. V-Model Otonom Sürüş Yazılım Mimarisi

Şekil 1'de açıklanabilir yapay zeka katmanlı V-model otonom sürüş yazılım mimarisi verilmiştir. Beş temel katmandan oluşan yazılım mimarisinin her katmanında algılama ve eyleyici (actuator) sınıfından birer modül hiyerarşik seviyeye göre konumlandırılmıştır. En altta yer alan İnsan Makine Arayüzü ve Gözlem modülü ise kullanıcı ile etkileşim ve sistemin minimum hata ile çalışmasından sorumludur. Katmanlar arası iletişim ardışıl olarak tanımlanmıştır ancak İnsan Makine Arayüzü ve Gözlem Modülü ise diğer modüllerden farklı olarak her katmandaki modüllerle iletişim halindedir. Otomotiv Mühendisleri Derneği (SAE) tarafından tanımlanan SAE J3016 standartında belirtilen üç temel fonksiyon grubu referans alındığında katmanlar aşağıdaki gibi gruplandırılabilir:

- **Operasyonel:** Sensörler - Eyleyiciler, Sensör Arayüzü - Eyleyici Kontrolü Katmanları
- **Taktiksel:** Algılama - Alt Seviye Kontrolör, Çevre Modeli - Üst Seviye Kontrolör Katmanları
- **Stratejik:** Açıklayabilir Yapay Zeka (AYZ) Yorumlayıcı - AYZ Planlayıcı Katmanı

Operasyonel modüllerde yüksek örneklemme zamanı gerekirken, taktiksel ve stratejik modüllerde nispeten daha düşük örneklemme zamanı kullanılabilir. Operasyonel modüllerde işlem yükü daha az iken stratejik modüllerde işlem yükü ve karmaşıklık artmaktadır.



Şekil 1: Otonom sürüş yazılım mimarisi.

3.1. Sensörler ve Eyleyiciler Katmanı

Bu katman, çevre ve araç platformu arasındaki arayüzü temsil eder. Günümüzde otonom araçlar, çevrelerindeki nesnelere algılamak ve girdi verisi oluşturmak için gelişmiş kameralar, ultrasonik sensörler, RADAR ve LIDAR gibi farklı türdeki sensörler

ile birlikte veri alımı (örneğin, V2X haberleşmesi) bileşenlerinden yararlanmaktadır.

Sensörler modülü, araca entegre sensörlerden ve dış veri kaynaklarının elde edilen ham bilgilerin toplanması ve yayılması görevini üstlenir. Bu modülün olası çıktıları arasında araca entegre sensörlerden alınan kamera görüntüleri ve nokta bulutunun yanı sıra radyo sinyalleri tabanlı V2X haberleşmesi yoluyla gelen araç ve trafik bilgisi gibi veriler de bulunabilir. V2X verisi, 3GPP c-V2X ve IEEE-802.11p standartları tabanlı CAM, CPM, DENM vb. mesaj tipleri olarak iletilir ve otonom araç sistemlerine yüksek veri alışverişi ve güncel bilgilere erişim sağlama avantajı sunarak güvenliğini artırır ve sürüş deneyimini iyileştirir.

Eyleme geçme tarafında, eyleyiciler modülünde, bu mesajlar içeriklerine bağlı olarak bazı ön işleme süreçlerinden geçirilip anlamlı hale getirildikten sonra ilgili modüllere (örneğin, algılama vb.) yönlendirilerek çıktı verisi olarak işlenir. Böylelikle kontrol sinyalleri, fiziksel kontrol edilebilir elemanlar olan gaz, fren ve direksiyon gibi eyleyicilere iletilir.

3.2. Sensör Arayüzü ve Eyleyici Kontrolü Katmanı

Sensör arayüzü modülünde, bir önceki katmandan alınan ham veriler çözülür, kontrol edilir, gruplanır ve ilgili veriler ilgili sensör arayüzlerine dağıtılır. Ayrıca bu katmanda birim ve koordinat dönüşümleri yapılır.

Eyleyici kontrolünde amaç eyleyicilerin planlanan durumu ile anlık durumu arasındaki farkı mümkün olduğunca düşük tutmaktır. Bu doğrultuda eyleyici elemanların kontrolcülerini gecikme, kavis sınırı, kavis oranı sınırı ve en yüksek ivmelenme ve yavaşlama (deceleration) değerleri gibi parametreler göz önünde bulundurularak tasarlanır.

Bu katmanda gürbüz kontrolcü algoritmaları tasarlanmalı ve ayrıntılı olarak test edilmelidir. Değerlendirme, algılayıcı çıktıları ile aracın planlanan durumu kıyaslanarak yapılabilir.

3.3. Algılama ve Alt Seviye Kontrol Katmanı

Algılama modülü filtreleme, takip etme ve özellik çıkarma (feature extraction) olmak üzere 3 ana adımdan oluşur. Filtreleme adımı, gürültüyü gidermek ve güvenilir veri elde etmek amacıyla uygulanırken takip etme adımı, takip edilen nesnenin hareketini takip eder ve nesnenin son durumunu günceller. Takip etmenin amacı, birden fazla kaynağı birleştirmeden her bir sensör verisinin Alt Seviyeli semantik yorumunu gerçekleştirmektir; örneğin kamera ve LIDAR gibi sensörlerden gelen veriler, nesnelere algılar ve takip etme algoritmalarıyla nesnelere hareketini analiz eder. Böylelikle nesnelere hız, yön ve pozisyon bilgileri güncellenir ve bu bilgiler takip edilmekte olan nesnenin gelecek bilgileri hakkında tahminde bulunulabilmesini sağlar. Algılama modülü, işaret tanıma, statik ve dinamik nesne algılama, şerit algılama, kaldırım kenarı algılama ve daha fazlası gibi sensör çeşitliliğine dayalı ek bileşenler de içerebilir. Özellikle çıkarma adımı, algılanan verilerden anlamlı bilgilerin çıkarılması ve çevre hakkında doğru yorumlamaların yapılması ile ilişkilidir. Yorumlanan veriler, sürücüye veya sistemi geliştiren kişiye sunulurken kullanılır. Sensörlerden gelen farklı nitelikli veriler işlenip dönüştürülür ve görsel veya anlamsal bir bilgi yaratılıp ilgili kişiye algılanan nesne açıklanır. Örneğin algılanan nesnenin ne olduğunun yorumlanması, trafik işaretlerinin an-

laşılması ve yol sınırlarının tespit edilmesi gibi bilgiler sunulmaktadır. Aracı kullanan sürücüye bu bilgiler gerçek zamanlı olarak sunulur, sistemi geliştiren kişiye yönelik ve daha gelişmiş ve yüksek hesaplama gücü ihtiyacı duyan bilgiler ise daha sonra kullanılmak üzere kayıt altına alınır ve işlenir. Bu adımda çeşitli yapay zeka algoritmalarından yararlanılarak nesnelere özellikleri tespit edilir ve çevre ile olan ilişkisi anlamlandırılır. Böylelikle, sürüş kararı ve planlama adımı için Üst Seviyeli bir çevre temsili oluşur.

Alt Seviye kontrol modülü boylamsal ve yanal kontrol olmak üzere iki bileşenden oluşmaktadır. boylamsal kontrolcü ve yanal kontrolcü için sırasıyla "Akıllı Sürücü Modeli" (Intelligent Driver Model, IDM)[20] ve "Şerit Değişiminden Kaynaklanan Toplam Frenlemeyi Minimizasyon" (Minimization Overall Braking Induced by Lane-change , MOBIL) [21] modelleri örnek gösterilebilir. Alt seviye kontrolcü modülünün temel görevi, işletme sınırları içinde istenen kontrol sinyallerini üretmektir. Kısaca, boylamsal kontrol hızlanma ve frenleme sistemlerini, yanal kontrol ise direksiyon açısını ayarlayan sistemleri kontrol eder. Bu iki modülün aynı katmanda değerlendirilmesinin nedeni, alt seviye kontrolün temel olarak ego aracının ve çevresinin durum bilgilerini kullanmasıdır ve algılama modülünün çıktılarının alt seviye kontrol modülünün çıktısını test etmek için yeterli olduğu söylenebilir.

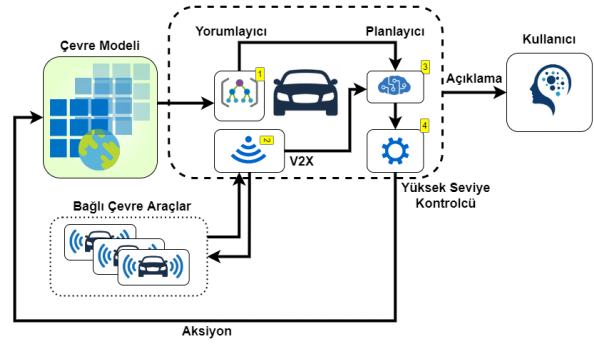
3.4. Çevre Modeli ve Üst Seviye Kontrol Katmanı

Çevre modeli katmanı, çeşitli algılama kaynaklarından gelen bilgileri birleştirerek aracın bulunduğu çevrenin temsili oluşturur. Bu katmanda algılanan nesnelere, şeritler, trafik akışı, bireysel izleyici çıktılar, harita bilgisinden ham verileri içeren girdiler toplanıp kullanılarak ortamın anlaşılması sağlanır. Bu veriler lokalizasyon ve sensör füzyonu olmak üzere iki ayrı amaç için kullanılır. Şeritlere veya yola göre bölünmüş bir lokalizasyon ile takip edilen şeritlerin, haritanın veya global yön bulma uydusu sisteminin verileri eklenir. Sensör füzyonu adımı, sürülebilir alanın yapısı ve araç durumu tahmin algoritmalarını içerir ve araç çevresindeki tüm nesnelere olasılık tabanlı algoritmalarla kullanılır, karakteristikleri birleştirilir ve çevre modeli çıkartılır.

Üst Seviye kontrol modülü planlayıcı modülünün çıktıları olan şerit değişikliği, şerit takibi gibi manevra kararları ile çevre modelinden gelen sürülebilir alan, şerit çizgileri ve trafik işaretleri gibi çevresel faktörlere göre aracın takip edeceği yörüngeyi lokal olarak hesaplar. Lokal koordinat eksenine göre belirli bir aralık için hesaplanan yörünge aracın konumu, oryantasyonu, hız ve ivmelenme değerlerini içermektedir. Aynı katmanda yer alan çevre modelinin çıktıları Üst Seviye kontrol modülünün çıktılarının performans değerlendirmesi için uygun ve özet veri içermekte ve V-döngüsü içerisinde bütünlük testlerinin yapılmasını kolaylaştırmaktadır. Ayrıca ilgili katmanda kurulan bu veri ilişkisi gerçek zamanlı sürüşte hataların ve sorunların daha erken yakalanabilmesine yardımcı olur.

3.5. AYZ Yorumlayıcı ve AYZ Planlayıcı Katmanı

AYZ Yorumlayıcı ve AYZ Planlayıcı Katmanı otonom araç mimarisi için oldukça önem taşımaktadır. Şekil 2'de, aracın çevre modelinden aldığı bilgileri işleyiş süreci, kullanıcı ve çevre araçlara işlenen bilginin aktarımı gösterilmektedir.



Şekil 2: AYZ yorumlayıcı ve planlayıcı katmanı.

Yorumlayıcı, çevre modeli katmanında bulunan trafik akışı, nesne ilişkileri ve trafik işaretleri gibi karakteristiklerini, geçmiş, şimdiki zamanın verilerini kullanarak gelecek zamana yönelik çevrenin semantik bir anlayışına çevirmekten sorumludur. Yorumlayıcı modülü, çevreyi doğru yorumlamak için yolun yasal kuralları ve sokak topolojisi ile dinamik ortam koşullarını soyut bir şekilde kullanır. Buna ek olarak gelecek tahminlerini yapabilmek için geçmiş trafik koşullarını da hesaba katar. Böylece dinamik haldeki her nesnenin gelecekteki konumları, tüm araçları açısından tehlike arz edecek durumları veya nesnelere sürücüye hızla aktarır hem optimum sürüş olanağı sağlar, ayrıca olası tehlikeleri önlemek üzere sürücüye uyarır. Gelecekteki durum tahminleri, davranış üretimi için risk değerlendirmesi getiren olasılıksal bir modelle birlikte gelir ve çıktılarını Planlama modülüne sağlar. Sistemin tanımlama ve yorumlamasına en çok katkı veren nesnelere çevre modelinde olduğu gibi sürücü veya geliştiricilere odak haritası olarak sunulur.

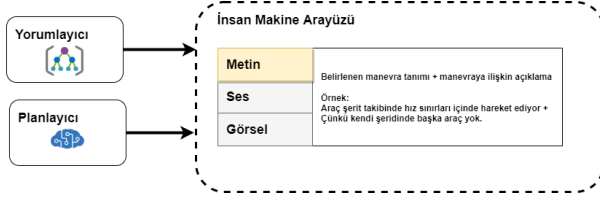
Planlayıcı, diğer bir adıyla karar verici, çevre modelinin yorumlamasını kullanarak olası davranış seçenekleri arasından en iyi davranışı seçer. Planlayıcının, son yıllarda pekiştirmeli öğrenme algoritmaları kullanılarak oluşturulması popüler hale gelmiştir. Pekiştirmeli öğrenme algoritması otonom araç için optimal davranışı oluşturur. Çevreyi ve sistem yeteneklerini anlayarak en iyi davranışı seçen planlayıcının, seçimlerini belirli sistem özelliklerine göre alması beklenir. Bu nedenle, araştırmacılar çalışmalarını açıklanabilir pekiştirmeli öğrenme algoritmaları üzerine yoğunlaştırmıştır [17], [22], [23]. Otonom araçlar için sistem dinamiklerini ve gözlemlerini açıklayan planlayıcılar karar mekanizmasını daha güvenilir hale getirmiştir [24].

Planlayıcı, çevre dinamiklerinin yorumlaması ve gözlemlerine dayanan kararını, bir sonraki katman olan üst seviye kontrolcüsüne yönlendiren bir gözetmen gibi davranır. Karar mekanizması, kurallara dayalı ve olasılığa dayalı durum geçiş yapısı içerisinde değerlendirilebilir. Modele bağlılık ve modelden bağımsızlık yapısına göre tüm davranışlar, önceliklerine ve teknik yönlerine dayanarak kategorilere ayrılabilir. Alınan kararlar ölçeklenebilir ve güvenilir özellikte uygulanabilir olmalıdır. Bununla birlikte, planlama modülü, karşılaşılan senaryoda uygun bir davranış olmadığında, sürüşü güvenli bir şekilde sürdürmek için koruma veya güvenli bir alana geçme gibi yeni seçenekleri aramak için çözüm alanını keşfeder.

Özetlemek gerekirse, bu katman, çevrenin anlamsal anlayışını, olasılıksal gelecek tahminini, davranış oluşturma ve davranış seçimi bileşenlerini içerir. Anlamsal anlayış ve olasılıksal gelecek tahmini bileşenleri yorumlayıcı modülünde ele alınır ve bu işlevlerin çıktılarını planlama modülüne sağlar. Davranış oluşturma ve davranış seçimi bileşenleri ise, uygun davranışları gerçekleştirmek için anlamsal bilgileri kullanarak planlama modülünde kullanılır.

3.6. İnsan Makine Arayüzü ve Gözlem Modülü

İnsan Makine Arayüzü ve gözlem modülü mimarinin her katmanı ile haberleşebilmekte ve hem algılama hem de eyleyici işlemlerini içermektedir. İnsan Makine Arayüzü otonom araç ve insanlar arasındaki iletişim için kullanılan donanım ve yazılım bileşenlerinden oluşur. AYZ yorumlayıcı ve planlayıcının çıktıları kullanıcıya literatürde yer alan mevcut doğal dil işleme yöntemleri ile metinsel ve işitsel olarak aktarılabildiği gibi, farklı görsel ifadelerle de aktarılabılır. Şekil 3'te AYZ yorumlayıcı ve planlayıcı çıktılarının arayüzte gösterimine ilişkin kavramsal tasarım verilmiştir. Gözlem birimi ise mimarideki her bir modülün fonksiyonel güvenliğini sağlayan hata tespiti ve yönetimi bileşenlerinden oluşur. Hataya düşen modüllerin devreden çıkarılması ve yeniden devreye alınması süreçlerini yönetir. Ve bu süreçleri yönetirken hataları yine AYZ Yorumlayıcı ve Planlayıcı katmanındaki çıktılarla ilişkilendirilerek anlaşılabilir bir yapıda kayıt eder.



Şekil 3: İnsan-makine arayüzü açıklanabilirlik kavramsal tasarımı.

4. Sonuç ve Öneriler

Bu çalışmada yazılım dünyasında yoğunlukla kullanılan V-model mimariye şekilsel ve işlevsel benzerlikleri olan bir otonom sürüş yazılım mimarisi önerilmiş ve mimarinin açıklanabilir, şeffaf, güvenilir ve otomotiv geliştirme standartlarına uyumlu olmasını sağlayacak yapıda olması hedeflenmiştir. Otonom sürüş yazılımlarında kullanılan yapay zeka ve derin öğrenme algoritmalarının açıklanabilir olmasına yönelik yapılan önemli çalışmalara değinilmiş ve bu yöntemlerin yazılım mimarisindeki yeri ve diğer katmanlarla olan ilişkisi detaylı bir şekilde sunulmuştur. Gelişmiş sensörler ve haberleşme yöntemleri sayesinde çevredeki nesnelere gelen ham bilgilerin V-model mimaride Sensör Arayüzü modülünden AYZ Yorumlayıcı ve Planlayıcı katmanına kadar akışı ele alınmıştır. AYZ katmanında çevre modelinden gelen girişlere göre en iyi davranış belirleyen ve otonom aracın almış olduğu karara ilişkin açıklamaları İnsan Makine Arayüzü ve Gözlem modülüne aktarabilen

bir yapı önerilmiştir.

Çalışmada sunulan mimari, gelecekte yaygınlaşması beklenen otonom araçlara yönelik tanımlanan isterleri karşılayabilecek, farklı donanımdaki araçlara aktarılabilen esnek ve kapsayıcı bir yapı olarak kurgulanmıştır. Önümüzdeki yıllarda özellikle mimarinin AYZ katmanındaki geliştirmelerin tamamlanması ve otonom sürüş yazılımının simülasyon ve gerçek ortamda test edilmesi hedeflenmektedir.

5. Kaynakça

- [1] Robert N Charette, "How software is eating the car," *IEEE Spectrum*, June, 2021.
- [2] Alexandru Constantin Serban, Erik Poll, and Joost Visser, "A standard driven software architecture for fully autonomous vehicles," in *2018 IEEE International Conference on Software Architecture Companion (ICSA-C)*. IEEE, 2018, pp. 120–127.
- [3] Wenhao Zong, Changzhu Zhang, Zhuping Wang, Jin Zhu, and Qijun Chen, "Architecture design and implementation of an autonomous vehicle," *IEEE access*, vol. 6, pp. 21956–21970, 2018.
- [4] Sirin Akkaya, Yesim Gurbuz, Mertcan G Zile, Engin Baglayici, Haydar A Seker, and Ahmetcan Erdogan, "A modular five-layered v-shaped architecture for autonomous vehicles," in *2019 11th International Conference on Electrical and Electronics Engineering (ELECO)*. IEEE, 2019, pp. 850–854.
- [5] David Gunning, Mark Stefik, Jaesik Choi, Timothy Miller, Simone Stumpf, and Guang-Zhong Yang, "Xai—explainable artificial intelligence," *Science robotics*, vol. 4, no. 37, pp. eaay7120, 2019.
- [6] Finale Doshi-Velez and Been Kim, "Towards a rigorous science of interpretable machine learning," *arXiv preprint arXiv:1702.08608*, 2017.
- [7] Yiran Xu, Xiaoyin Yang, Lihang Gong, Hsuan-Chu Lin, Tz-Ying Wu, Yunsheng Li, and Nuno Vasconcelos, "Explainable object-induced action decision for autonomous vehicles," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020, pp. 9523–9532.
- [8] Zhengming Zhang, Renran Tian, Rini Sherony, Joshua Domeyer, and Zhengming Ding, "Attention-based interrelation modeling for explainable automated driving," *IEEE Transactions on Intelligent Vehicles*, vol. 8, no. 2, pp. 1564–1573, 2022.
- [9] Jinkyu Kim and John Canny, "Interpretable learning for self-driving cars by visualizing causal attention," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 2942–2950.
- [10] Luca Cultrera, Lorenzo Seidenari, Federico Becattini, Pietro Pala, and Alberto Del Bimbo, "Explaining Autonomous Driving by Learning End-to-End Visual Attention," *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, June 2020.

- [11] Mariusz Bojarski, Anna Choromanska, Krzysztof Choromanski, Bernhard Firner, Larry Jackel, Urs Muller, and Karol Zieba, "Visualbackprop: visualizing cnns for autonomous driving," *arXiv preprint arXiv:1611.05418*, vol. 2, 2016.
- [12] T Hong, Marilyn Abrams, Tommy Chang, and Michael Shneier, "An intelligent world model for autonomous off-road driving," *Computer Vision and Image Understanding*, vol. 80, no. 11, pp. 1–16, 2000.
- [13] Zhongwei Yu, Jingqing Ruan, and Dengpeng Xing, "Explainable reinforcement learning via a causal world model," *arXiv preprint arXiv:2305.02749*, 2023.
- [14] Mark Stefik, Michael Youngblood, Peter Pirolli, Christian Lebiere, Robert Thomson, Robert Price, Lester D Nelson, Robert Krivacic, Jacob Le, Konstantinos Mitsopoulos, et al., "Explaining autonomous drones: An xai journey," *Applied AI Letters*, vol. 2, no. 4, pp. e54, 2021.
- [15] Giuseppe Cascarino, Mirko Moscatelli, and Fabio Parlapano, "Explainable artificial intelligence: interpreting default forecasting models based on machine learning," *Bank of Italy Occasional Paper*, , no. 674, 2022.
- [16] Roman Liessner, Jan Dohmen, and Marco A. Wiering, "Explainable Reinforcement Learning for Longitudinal Control.," *International Conference on Agents and Artificial Intelligence*, 2021.
- [17] Jeff Druce, Michael Harradon, and James Tittle, "Explainable artificial intelligence (xai) for increasing user trust in deep reinforcement learning driven autonomous systems," 2021.
- [18] Gaith Rjoub, Jamal Bentahar, and Omar Abdel Wahab, "Explainable AI-based Federated Deep Reinforcement Learning for Trusted Autonomous Driving," *International Conference on Wireless Communications and Mobile Computing*, May 2022.
- [19] Ignacio Llatser, Thomas Michalke, Maxim Dolgov, Florian Wildschütte, and Hendrik Fuchs, "Cooperative automated driving use cases for 5g v2x communication," in *2019 IEEE 2nd 5G World Forum (5GWF)*, 2019, pp. 120–125.
- [20] Martin Treiber, Ansgar Hennecke, and Dirk Helbing, "Congested traffic states in empirical observations and microscopic simulations," *Physical review E*, vol. 62, no. 2, pp. 1805, 2000.
- [21] Arne Kesting, Martin Treiber, and Dirk Helbing, "General lane-changing model mobil for car-following models," *Transportation Research Record*, vol. 1999, no. 1, pp. 86–94, 2007.
- [22] Lindsay Wells and Tomasz Bednarz, "Explainable ai and reinforcement learning—a systematic review of current approaches and trends," *Frontiers in artificial intelligence*, vol. 4, pp. 550030, 2021.
- [23] Gaith Rjoub, Jamal Bentahar, and Omar Abdel Wahab, "Explainable ai-based federated deep reinforcement learning for trusted autonomous driving," in *2022 International Wireless Communications and Mobile Computing (IWCMC)*. IEEE, 2022, pp. 318–323.
- [24] AV Shreyas Madhav and Amit Kumar Tyagi, "Explainable artificial intelligence (xai): connecting artificial decision-making and human trust in autonomous vehicles," in *Proceedings of Third International Conference on Computing, Communications, and Cyber-Security: IC4S 2021*. Springer, 2022, pp. 123–136.