

(令和元年度戦略的基盤技術高度化支援事業)
「自律的自動運転の実現を支える人工知能搭載システムの安全性立証技術の研究開発」

研究開発成果報告

令和 2 年 3 月

事業管理機関：株式会社ヴィッツ
国立大学法人名古屋大学
主たる研究実施事業者：株式会社ヴィッツ
国立大学法人名古屋大学
株式会社アトリエ
アーク・システム・ソリューションズ株式会社

目次

1. 研究開発の概要	3
1.1. 研究開発の背景	3
1.2. 研究目標	4
1.3. 研究開発の取り組みの評価	5
2. 導入した技術、機器設備について	7
2.1. 導入した技術について	7
2.1.1. 認証機関 SGS ジャパンによる Safety-AI のレビュー	7
2.2. 導入した機器設備について	7
2.2.1. ZMP® PowerUnit Z	7
2.2.2. 仮想化検証システム「ViViD」オプション機能（7 種類）	7
3. 補助事業の具体的内容	8
3.1. 補助事業の具体的内容	8
3.1.1. 人工知能搭載システムを安全に活用するための対応（サブテーマ 1）	8
3.1.2. 人工知能搭載システムの潜在リスクへの対応（サブテーマ 2）	8
3.1.3. 人工知能搭載システムの開発コスト対応（サブテーマ 3）	9
3.2. 重点的に実施した事項	10
3.2.1. 人工知能搭載システムを安全に活用するための対応（サブテーマ 1）	10
3.2.2. 人工知能搭載システムの潜在リスクへの対応（サブテーマ 2）	11
3.2.3. 人工知能搭載システムの開発コスト対応（サブテーマ 3）	11
4. 補助事業の成果及びその効果	12
4.1. 人工知能搭載システムを安全に活用するための対応（サブテーマ 1）	12
4.1.1. Safety-II に基づく人工知能搭載システムの安全分析	12
4.1.2. 人工知能搭載システムの安全性説明手法の確立	21
4.1.3. 令和元年度の活動まとめ	29
4.2. 人工知能搭載システムの潜在リスクへの対応（サブテーマ 2）	31
4.2.1. 安全認証に必要な各種ドキュメントのテクニカルレポート取得（サブテーマ 2-2）	34
4.2.2. 人工知能の種別・安全度・安全対策方法の分類（サブテーマ 2-3）	41
4.2.3. 令和元年度の活動のまとめ	49
4.3. 人工知能搭載システムの開発コスト対応（サブテーマ 3）	49
4.3.1. 自動走行車両による実証実験	49
4.3.2. 「Trust AI Shield」	62
4.3.3. 「Fuzzing for AI」	74
5. 補助事業の成果に係る事業化展開について	86
5.1. 想定している具体的なユーザー、マーケット及び市場規模等に対する効果	86
5.2. 事業化見込み（目標となる時期・売上規模）	86
5.3. 事業化に至るまでの遂行方法や今後のスケジュール	87
5.4. 成果（試作品）の無償譲渡や無償貸与	88
6. 補助事業の成果に係る知的財産権等について	89
6.1. 知的財産権の出願及び取得並びに論文掲載の有無	89
6.2. ライセンス契約等による事業展開	89

1. 研究開発の概要

1.1. 研究開発の背景



本研究は、近年急速に普及拡大している人工知能 (Artificial Intelligent 通称 AI) を搭載した、自律的自動運転を含む組込み制御システムにおける安全性立証技術を実現する研究事業である。

人工知能は 1940 年代にニューラルネットワーク (神経接続網) をコンピュータ上で模擬する試みが端緒である。その後 1960 年代に思考論理を高度化した人工知能が発表され、人工知能の世界が開いた。2006 年にジェフリー・ヒントンによりオートエンコーダおよびディープ・ビリーフ・ネットワークが提案され、この研究がディープラーニング (深層学習) へとつながり、現在に至る人工知能ブームを牽引している。ディープラーニング以前と以後の大きな違いは、ディープラーニングは問題の解答を自動的に生成することができる点であり、その結果、「自ら学ぶことができる人工知能」の実現へ大きく前進している点である。

2016 年は人工知能を活用した技術が大きく展開されることになった。Google 社関連会社による「AlphaGo」がプロ棋士を相手に圧勝を収め、シンシナティ大学が開発した戦闘機用人工知能が元米軍パイロットとの模擬戦で勝利したりなど、特定の分野においては人間の能力を超える人工知能が登場している。

このような背景から、組込み制御システムにおいて人工知能搭載が加速されている。自律ロボットには人工知能の搭載が進んでおり、自動車メーカーは自律的自動運転車両の公道試験を実現するなど成果を上げつつある。

しかしながら自律制御における人工知能の活用では課題も出ている。2016 年 3 月に Google 社が開

発していた自動運転車両が交通事故を起こした。人工知能による周囲の障害物検知において後続車両の存在を考慮できなかったためである。この事故における主たる原因は人工知能搭載システムの認識技術に課題があることとされた。

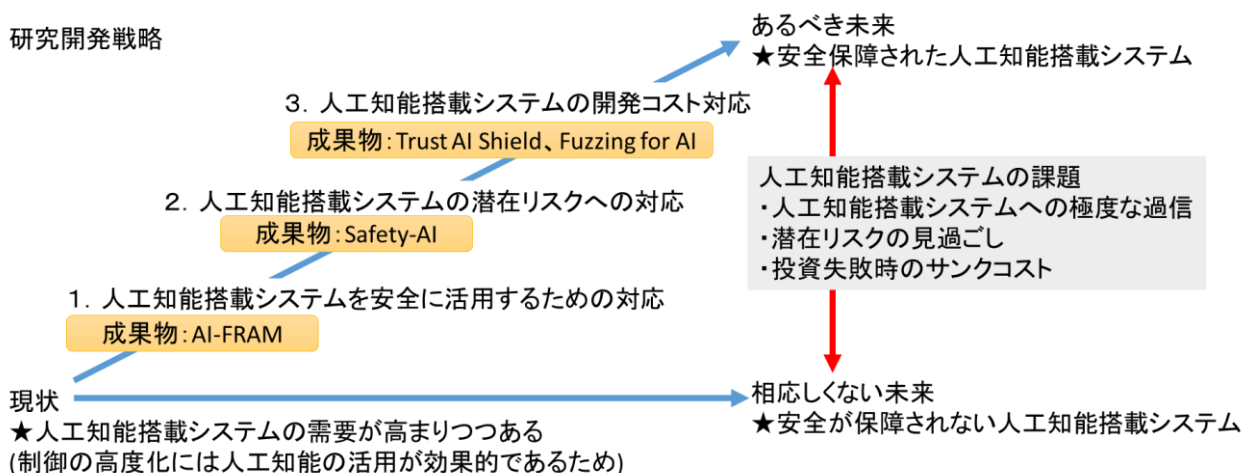
自動車メーカーの自動運転車両開発スケジュールでは、2019年に自動運転レベル2（高速道路等において運転者がハンドルを握った状態での自動運転）の製品をリリースし、2025年に自動運転レベル4（運転者不在での自動運転）の実現を計画しているが課題が山積している。克服が困難な課題としては、センサ技術・認識技術・判断技術・高度地図情報活用が挙げられ、これら重大課題について自動車メーカー各社やIT企業などが課題解決を急いでいる。従来の自動車制御技術、安全設計技術に加えて、変化していく走行環境への柔軟な対応は人工知能技術を活用しなければ実現できない。日本においては自動運転に人工知能を活用するため約1,096億円（2015年：EY総研調査データより）の研究費用を投じて対応している。また、人工知能関連産業市場予測では、運輸・輸送の自動化への期待が大きく、2030年には約30兆円市場（2015年：EY総研調査データより）へ成長するとの試算もある。

現在の自動車開発は安全設計が必要であり、自動車機能安全規格ISO26262の適合が製品に課せられている。しかし、人工知能搭載システムを含めた機能安全規格準拠の製品は前例が無い。そのため人工知能搭載システムによる自動運転車両の製品化は困難な状況にある。2017年には人工知能学会より「人工知能学会 倫理指針」が示され、人工知能が順守すべき事項が提案された。人工知能を安全に活用するための方針として国立研究開発法人産業技術総合研究所から「人工知能安全三方針」が提案され人工知能の安全について議論が活発化している。別の視点では人工知能そのものを安全とするのではなく、人工知能を含めシステム全体での安全を説明する新しい安全の概念が提案されるなど、人工知能を安全に活用するための活動がようやく進み始めた段階である。

現在のところ、自律的自動運転の実現を支える人工知能搭載システムの安全性立証技術は確立されておらず、川下企業においては人工知能搭載システムの安全性を立証するために100億キロ以上を走行させ、計測統計的なデータを用いて開発および試験を進めており、開発コストの試算ができていない状況である。

本研究では、人工知能搭載システムを安全に活用するための安全ガイドライン、安全分析手法、安全対策を実現し、人工知能搭載システムの安全性立証技術を確立する。

1.2. 研究目標



人工知能搭載システムの安全性立証技術を実現する。特に近年実現が要望されている自律的自動運

転において、人工知能搭載システムを安全開発できる実証例を示し、本研究成果の製品活用による更なる国際競争力の向上を実現する。

本研究で得られた成果は、現在、我が国の自動車メーカー各社が欧米諸国との激しい競争のなかで推進している自律的自動運転車両の開発において、実用化のための最大の課題である安全性立証の問題を解決し、国際競争力の強化に寄与するものである。

人工知能搭載システムを活用した自律的自動運転車両を、現在の自動車業界における基本的な安全概念として確立している国際規格 ISO 26262 の枠組みで実現する技術方策は、2017 年時点では世界的に見ても確立されておらず、各国の企業、機関とも課題を認識し検討を始めた段階である。本研究の実施グループである(株)ヴィッツ、(株)アトリエは、自動車分野における機能安全規格への対応に関して国内業界をリードする技術力を有しており、現時点から研究着手することで、欧米諸国の企業・機関に先駆けて課題解決の道筋を確立できる可能性は十分に高い。

国内自動車メーカーが本研究の成果を活用して自動運転車の実用化に成功し、世界の自動車業界における自律的自動運転車両の安全性立証技術を我が国主導で構築できたならば、圧倒的な国際競争力を確保し自律的自動運転車両の開発において技術的主導権を握ることも期待できる。

本研究のアドバイザー企業である世界トップの自動車メーカーの自動車生産台数は約 1,021 万台(2016 年度)であり、その生産台数の半数に自動運転オプション機能(20 万円/台)が搭載されたと仮定した場合、その売上効果は年間約 1 兆円 (1,000 万台×10 万円×1/2)と試算できる。

人工知能搭載システムの安全立証技術はすでに説明したとおり、自動車ばかりでなく他の制御機器に活用できる技術である。すなわち、来るべき IoT/CPS 社会において人類に有益な生活支援を行う知能化機械に波及する活用されるため、その波及分野・波及機器、さらには社会システムへの効果が期待できる。

本研究は人工知能搭載システムの安全性立証を主題としているため、人工知能を活用した、全産業分野および CPS/IoT 構成デバイス全てに当該研究成果の利用が期待できる。すなわち、次世代社会システムを構築する機器は少なからず他機器との連携やクラウド連携を行い、その連携は人工知能搭載システムによる制御ロジック機構を保有している。その人工知能搭載システムの安全性立証技術であるため、人工知能を安全に活用するための基盤技術として活用されることが想定される。

自動車以外の分野でも、人工知能搭載システムの安全性立証は重要な課題である。人と同じ空間で活動するサービスロボット、発電所・航空管制などの社会重要インフラなど、高度な知的判断能力と安全性の両立が求められる応用分野において、人工知能搭載システムが今後ますます広がっていくことは確実であるが、それらの分野全てにおいて研究活用が期待されるため、将来的な波及効果は計り知れない。

中部経済圏は世界有数の自動車メーカー、航空宇宙産業、産業機械・ロボットメーカーが多数存在し、これらの製品の高度化には人工知能搭載システムの活用が必須であるとともに、いずれの製品も高い安全性を求められる製品分野である。また、中部地区には次世代自動車産業地域産学官フォーラムを中心とした自動運転技術向上に関する活動の支援を行うなど、他地域と比較して、自動運転技術向上には力を入れており、自律的自動運転を支える人工知能搭載システムの活用とその安全性立証に関する研究は、当該地域の方針と合致している。

1.3. 研究開発の取り組みの評価

本研究事業においては、実施内容をサブテーマ 1～3 までの項目に分けて並行で実施し、研究を効率的に進めるために、サブテーマ毎のリーダーを据えて着実に研究が前進するための研究開発体制を整えている。

研究テーマ全体としては、本年度の大目標である、「人工知能を搭載したシステムの安全ガイドラ

インの完成」を中心におき、計画通りに研究事業として完了できた。

最も難易度が高いガイドラインの策定については、現行業界で活用がすすむ ISO26262 との融合が実現できたことが評価できる。

以下にサブテーマ毎の目標達成度を列挙する

サブテーマ１：人工知能搭載システムを安全に活用するための対応(進捗率 100%)

年度目標：人工知能の特性を明確にし、安全分析のための基本調査を完了

Safety - II に基づく人工知能搭載システムの安全分析により、人工知能搭載システムを含めたシステム全体による安全性は担保できうることを調査している。様々な論文と文献を調査し、対象システムを定め具体的な分析方法を検討している。調査の結果から、安全分析を実施するための手順を明確にした、手順書の作成に着手しており、想定通り、もしくはそれ以上の成果を上げつつある。

サブテーマ２：人工知能搭載システムの潜在リスクへの対応(進捗率 100%)

年度目標：人工知能搭載システムの安全を立証するための手法および帳票類の規定完了

平成 30 年度の調査・検討結果を踏まえ、安全を立証するための手法および帳票類の規定として、人工知能搭載システムを安全に開発するためのガイドラインを「Safety-AI」として作成を完了する。人工知能の種別や利用方法などをまとめた分類を作成し、それぞれのアーキテクチャから考えられる安全度、安全対応策などをまとめる。

さらに、この規定の活用が安全上技術的に適合していることについて、認証機関の確認を受け、「テクニカルレポート」を作成する。

サブテーマ３：人工知能搭載システムの開発コスト対応(進捗率 100%)

年度目標：人工知能搭載システムの開発において利用可能なソフトウェア部品の分析および設計の実施

人工知能搭載システムの機能不全検知と防衛機能「Trust AI Shield」の開発を行うために基本調査を実施している。また、人工知能搭載システムの動作検証「Fuzzing for AI」の開発のための人工知能搭載システムの機能不全を未然に解析できることを調査している。

調査の結果、具体的な設計工程を開始させており、十分な進捗があると考えられる。

【達成度に関する評価】

今年度計画したサブテーマ①、サブテーマ②、サブテーマ③は目的を達しており、研究活動の目標は達成されたと評価できる。

また、本事業を進めていくなかで、外部からの関心の高さが伺え、プロジェクトの進捗に興味を示す組織が複数あり、英国 Assuring Autonomy International Programme における共同研究プロジェクト TIGARS、AI プロダクト品質保証コンソーシアム QA4AI への委員として活動している。

令和元年度において平成 30 年度の成果物の販売を 3 件実現しており、事業化へと進んでいる。

2. 導入した技術、機器設備について

2.1. 導入した技術について

2.1.1. 認証機関 SGS ジャパンによる Safety-AI のレビュー

2019 年 11 月 18～19 日、2020 年 1 月 29 日に、認証機関 SGS ジャパンによる Safety-AI のレビューを受けた。詳細についてはサブテーマ 2 の報告にて記載する。

2.2. 導入した機器設備について

2.2.1. ZMP® PowerUnit Z

サブテーマ 1 および 3 で活用。

平成 29 年度に導入した NVIDIA 社 Drive PX2 をゴルフカート上で動作させるためのバッテリーユニット。PowerUnit Z を用いて実証実験の実現について検討を実施した。

2.2.2. 仮想化検証システム「ViViD」オプション機能（7 種類）

サブテーマ 1 および 3 で活用。

平成 29 年度に導入した ViViD の機能拡張オプション。安全アセスメントを実施するための各種パッケージを合計 7 種類導入し、安全アセスメントの実現について検討を実施した。

3. 補助事業の具体的内容

3.1. 補助事業の具体的内容

3.1.1. 人工知能搭載システムを安全に活用するための対応（サブテーマ 1）

令和元年度（以下本年度）のサブテーマ 1 の活動は、平成 30 年度（以下昨年度）に実施した「人工知能搭載システムの安全性立証を実現するための安全分析手法の確立」を基に人工知能搭載システムの安全分析の完了のための研究として以下の細目について活動した。

3.1.1.1. Safety-II に基づく人工知能搭載システムの安全分析（サブテーマ 1-1）

本研究の出発点は、「人工知能を搭載した自律運転システムを利用者は受け入れることができるのか」であり、この社会的受容性を満たす上で大きな課題となるのが人工知能特有の課題を抱えたシステムを安全に運用ができるのかという点である。ここで本研究が指す「安全」とは国際基本安全規格（ISO/IEC GUIDE 51:2014）に定義されている、「許容できないリスクがないこと」と定義している。

サブテーマ 1 では昨年度の研究から提言した「安全分析手法の確立」から、提言した手法を基にした「人工知能搭載システムにおける安全分析を完了」を本年度目標として活動した。

3.1.1.2. 人工知能搭載システムの安全性説明手法の確立（サブテーマ 1-2）

本研究を事業化しサービスを立ち上げる場合には、自治体や自動車メーカー、サプライヤーやサービス事業者など様々なステークホルダーが関わる事になる。一方でステークホルダーによってシステムに関する知識や理解には差があり、自動運転システムという新しい技術の導入には漠然とした不安がある事も考えられる。これに対応するためにステークホルダーに合わせた内容をより伝えやすくするための説明手法を確立させ、人工知能搭載システムの安全性について説明をする必要がある。そこで、サブテーマ 1 では「人工知能搭載システムの安全性説明手法の確立」を本年度目標として活動した。

3.1.2. 人工知能搭載システムの潜在リスクへの対応（サブテーマ 2）

令和元年度においては、人工知能搭載システムを安全に開発するためのガイドライン「Safety-AI」を更新し、妥当性確認を認証機関に依頼し、「テクニカルレポート」を取得することを研究目標として、以下の細目について活動を実施した。

3.1.2.1. 安全認証に必要な各種ドキュメントのテクニカルレポート取得（サブテーマ 2-2）

自動車の機能安全規格 ISO 26262 では、人工知能搭載システムの活用を認めてはいない。そのため人工知能搭載システムの自動車への導入と自動車製品化に必要な機能安全適応は容易に実現できない。国際規格においては、利用を認めていない技術を活用するにはその技術の利用が妥当かつ適切な対処が施されていることを規格が認める方法で説明する必要がある。説明を実施するためには、安全コンセプト、安全マニュアル、開発プロセスなど各種のドキュメントを作成し、適切な認証機関に確認してもらうことが重要である。

平成 29 年度は、認証機関による内容検証により、技術的に適合していることを示す「テクニカルレポート」を取得するための準備を実施した。

平成 30 年度においては、平成 29 年度の調査・検討結果を踏まえ、さらにサブテーマ 2-1 で作成した Safety-AI に従って安全コンセプトを構築し、認証機関による内容検証を受けた。人工知能周辺の安全設計については概ね問題ないが、人工知能搭載システムが安全かを評価するには、システム全体の定義やハザード分析&リスクアセスメントの厳密化が必要との結果に至った。

令和元年度においては、平成30年度の認証機関からの助言を踏まえ、自動運転レベル4システムの安全コンセプトを再構築し、認証機関による内容妥当性確認を受け、「テクニカルレポート」を取得した。

3.1.2.2. 人工知能の種別・安全度・安全対策方法の分類（サブテーマ2-3）

平成30年度においては、安全分析を容易にするために、人工知能の種別や利用方法などをまとめた分類をまとめた。また、人工知能の特性・特徴、それぞれのアーキテクチャから考えられる安全対応策などをまとめた。これらの知見をまとめ、Safety-AI に記載した。

令和元年度においては、人工知能搭載システムまで考慮範囲を広げ、システム定義や網羅的なハザード分析&リスクアセスメントの厳密化を行うための観点の分類整理を行い、Safety-AI に記載した。また、各種アーキテクチャに対する安全対策方法について詳細化を行った。

3.1.3. 人工知能搭載システムの開発コスト対応（サブテーマ3）

3.1.3.1. 自動走行車両による実証実験（サブテーマ3-1）

平成30年度に引き続き、「Trust AI Shield」の搭載と「Fuzzing for AI」による検証を例示するための研究を実施した。

今年度は「Trust AI Shield」および「Fuzzing for AI」の再設計と実装を完了し、シミュレータ環境および自動運転車両を用いた実機環境の両方で実証実験を実施した。実施した内容と結果を以下に示す。

（1）実証実験 ACC システム

実証実験用の ACC (Adaptive Cruise Control) システムを再設計および実装した。仮想環境には「ViViD (Virtualized Verification automatic Driving)」(以下、ViViD とする)を用いて動作確認を行ったのちに、自動走行車両を用いた実機環境での実証実験を実施した。

実証実験は名古屋大学のシャーシダイナモにて、動作確認テストを34件、走行テストを39件実施した。これらのテスト項目をすべてパスすることで、ACC システムが意図通りに動作することを確認した。シャーシダイナモ施設は屋内施設であるが、本来の想定環境は屋外であるため、屋外でのセンサ認識結果に問題がないことを確認した。

（2）「Trust AI Shield」

「Trust AI Shield」の判断に用いるために、AI に影響を与える可能性があるファクターリストを作成した。このファクターによってAI の誤検出・誤判断が生じ、AI 搭載システムの機能不全が引き起こされる可能性がある。そのため評価の際にファクターの影響を考慮しておく必要があり、このリストはAI 搭載システムの評価項目として転用できると考えている。本実証実験システムでは、ファクターリストのうち、天候・時刻を利用した「Trust AI Shield」の設計および実装を行った。「Trust AI Shield」を使用することで、AI の出力結果に対して、信頼度を高めることができた。

「Trust AI Shield」の機能の一部として、走行時のログを保存する機構を実装した。ログを保存しておくことで、走行後に不具合の原因を解析することができるため、システムの保守に役立てることができる。

（3）「Fuzzing for AI」

ViViD を用いて自動で生成した機械学習アノテーション情報を正解データとし、物体検出 AI の出力が正しいかどうかを自動で検証するツールを設計および実装した。「Fuzzing for AI」にはノイズ付与機能を合わせて実装した。ノイズ付与機能により、テストデータのバリエーションを効率的に増やすことができる。加えて、ノイズの有無で物体検出 AI の精度に変化が生じるかを比較することで、AI がどのようなノイズに影響を受けるかを確認することができる。

ViViD と「Fuzzing for AI」を用いたテストを1024件実施した。これにより、ACC システムに用いる物体検出 AI がどのような状況で誤検出・誤判断しやすいかを、自動で効率的に検証することがで

きた。手作業で検証を実施した場合の作業効率は1時間当たり約35件であるが、本検証ツールを使用すると1時間当たり約3600件の検証が可能になる。「Fuzzing for AI」を使用することで、効率的で現実的なコストでの物体検出AIの検証を実現できた。

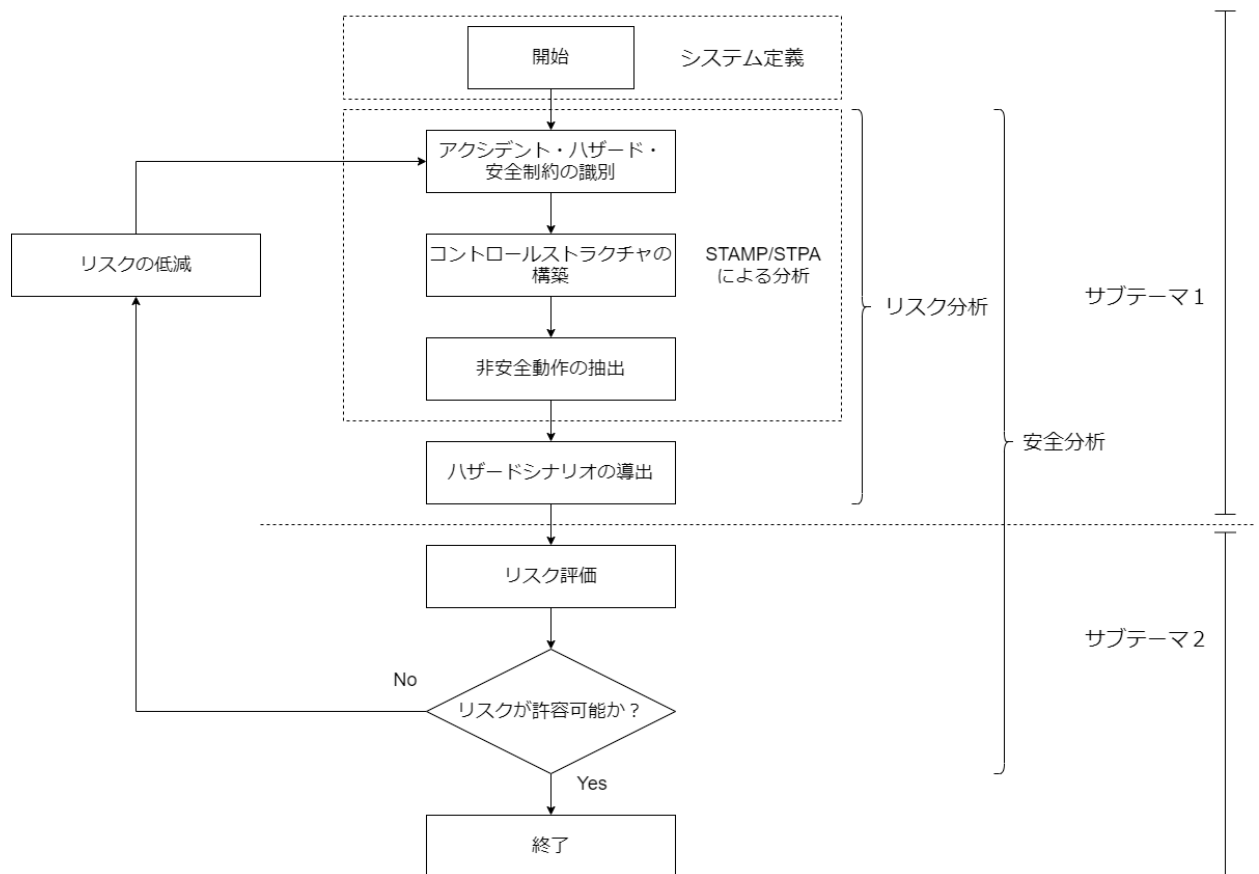
3.2. 重点的に実施した事項

3.2.1. 人工知能搭載システムを安全に活用するための対応(サブテーマ1)

本年度のサブテーマ1の活動は、サブテーマ2と共同で人工知能搭載システムの安全分析の完了を目標として安全分析を実施した。安全分析を網羅的に実施するには前提としてシステム定義が明確である必要があるという課題が昨年度の活動によって明確になっていた為、サービスレベルを想定した人工知能搭載MaaSシステムのシステム定義を実施した。

安全分析の目的は対象のシステムにおいて「許容できないリスクがないこと」を実現することである。以下に安全分析のプロセスを示す。

本研究での安全分析のプロセス



この安全分析のプロセスの内、サブテーマ1ではSTAMP/STPAを用いたハザード分析と、ハザードシナリオの導出を実施した。そのハザードシナリオに対するリスク評価とリスク低減の為の対策検討をサブテーマ2で実施した。ハザード分析手法には昨年度の活動から人工知能搭載システムに対して適用が可能な手法である事を示したSTAMP/STPAを選定した。

3.2.1.1. Safety-IIに基づく人工知能搭載システムの安全分析(サブテーマ1-1)

本年度の活動では、昨年度の車両単体を対象とした安全分析から拡大し、サービスを視野に入れた

MaaS システムを対象とした安全分析を実施した。MaaS とは Mobility as a Service の略称で移動サービスを意味する。今回は自動運転車両を主役として、インフラ機能で走行を補助する人工知能搭載 MaaS システムを定義し、そのシステムに対する安全分析を実施した結果から安全性を担保する為の考え方を示した。

3.2.1.2. 人工知能搭載システムの安全性説明手法の確立（サブテーマ 1-2）

本年度の活動では、人工知能搭載自動運転システムの事業化を提案する上で様々なステークホルダーに対して複雑なシステムや安全性への説明をより伝えやすくする手法が必要であるとして、図式を用いるビジュアル化による説明手法を提言した。ビジュアル化の手法として「俯瞰図」「イラスト」「3D シミュレータ」の3つを提言し、各手法の特徴や用いられた実例からより伝えやすい説明手法についての考え方を示した。

3.2.2. 人工知能搭載システムの潜在リスクへの対応（サブテーマ 2）

委員会におけるアドバイザー/オブザーバからの意見、認証機関のレビュー結果を踏まえ、人工知能搭載システムの安全に関して多角的な分析・検討を重ね、その知見をまとめてガイドライン「Safety-AI」第2版に更新した。

3.2.2.1. 安全認証に必要な各種ドキュメントのテクニカルレポート取得（サブテーマ 2-2）

具体的な人工知能搭載システムとして、限定区域内の自動運転レベル4システムを想定システムとした上で、サブテーマ1と合同でシステム定義ならびにハザード分析&リスクアセスメントを実施し、安全設計を検討し、安全コンセプト文書を作成した。さらに、認証機関による技術レビューを受け、人工知能搭載システムの安全設計に対する我々の取り組みの方向性は正しいことを確認できた。

3.2.2.2. 人工知能の種別・安全度・安全対策方法の分類（サブテーマ 2-3）

国際規格など整理された知見が一切発行されていない中、以下の知見を整理し、ガイドラインの充実を図った。

- ・システム定義や網羅的なハザード分析&リスクアセスメントに関する実施方法や観点
- ・人工知能搭載システムの安全論証手法の分類と具体的な手順
- ・人工知能搭載システムの安全説明力の高い開発プロセス

さらに、これらの知見について対外紹介活動を行った。具体的には、AI プロダクト品質保証コンソーシアム（QA4AI）や人工知能に関する標準化委員会である ISO/IEC JTC1 SC42 と連携し、セミナーによる技術紹介を実施し、高い関心・評価を得ることができた。

3.2.3. 人工知能搭載システムの開発コスト対応（サブテーマ 3）

3.2.3.1. 自動走行車両による実証実験

自動走行車両とシャーシダイナモを使って、実証実験を行った。

具体的なソフトウェアコンポーネントとして初年度から検討を続けてきた、「Trust AI Shield」の設計および実装を完了した。また、完成させるだけでなく、完成したコンポーネントを用いてテストを実施することで、AI を安全に活用するための対策を正常に運用または動作することを示した。

4. 補助事業の成果及びその効果

4.1. 人工知能搭載システムを安全に活用するための対応(サブテーマ 1)

4.1.1. Safety-II に基づく人工知能搭載システムの安全分析

4.1.1.1. 人工知能搭載 MaaS システムの定義

安全分析は使用条件や環境条件によって考慮すべきリスクが異なる。サブテーマ 1 の本年度の活動ではまず MaaS システムの定義を行い、どのようなシステム・制限事項・環境条件なのかを明確化した。以下にシステムの概要図を示す。

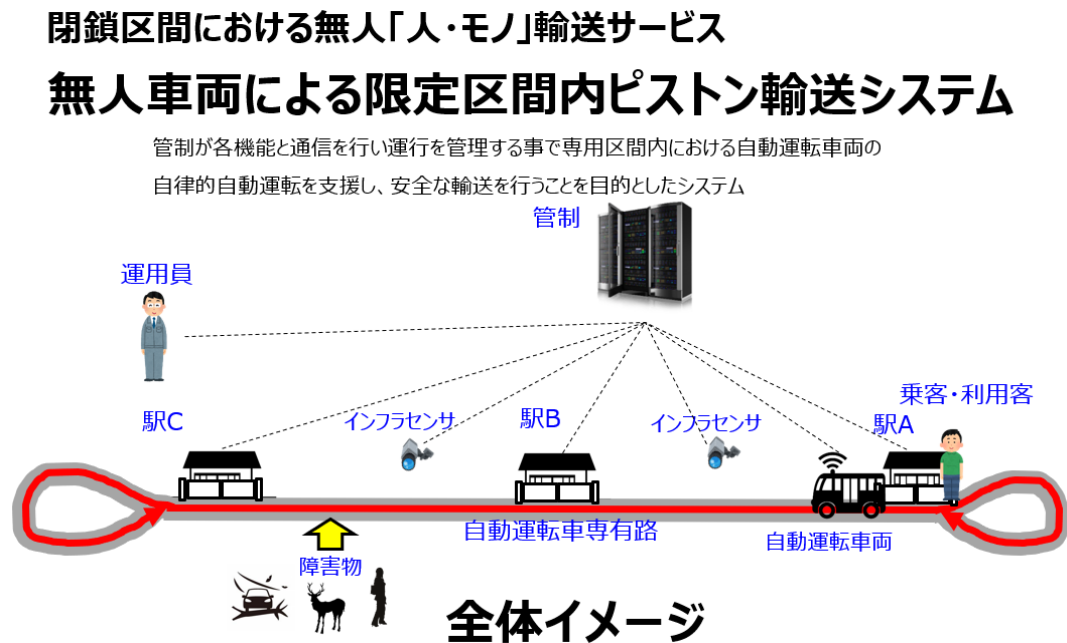
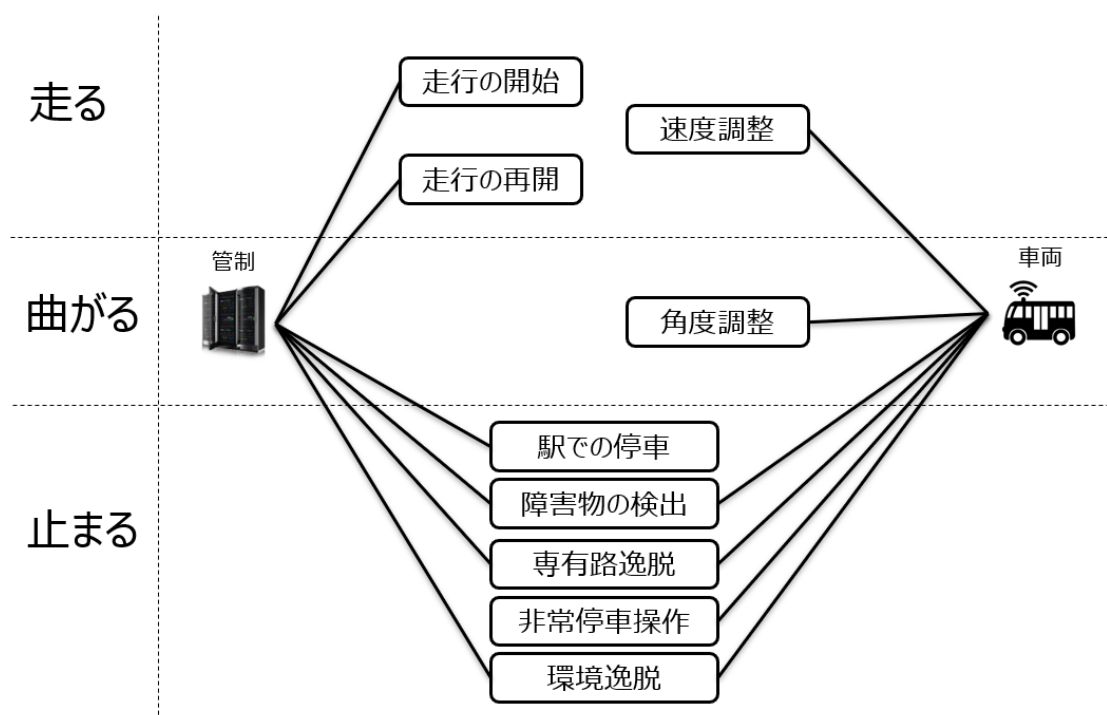


図 システムの概要図

本システムは自動運転車（以下、車両）を一般車・一般人の立ち入らない閉鎖区間内の自動運転車専用路（以下、専用路）で走行させ、利用者を輸送する輸送サービスシステムであり、管制によって制御される各機能の補助によって安全な走行を実現する事を目的とする。

本システムでの自動運転レベルは限定区間内でシステムが全てを操作するレベル 4 の自動運転システムによる走行を想定し、車両と管制で走行機能を分担して制御する事で人による運転操作は不要になっている。本システムにおける人工知能は車両が走行する際の白線検出や、障害物との衝突回避を行う為の障害物検出を行う画像認識が役割となる。以下に管制と車両の車両制御の分担を示す。



a. システムの登場人物

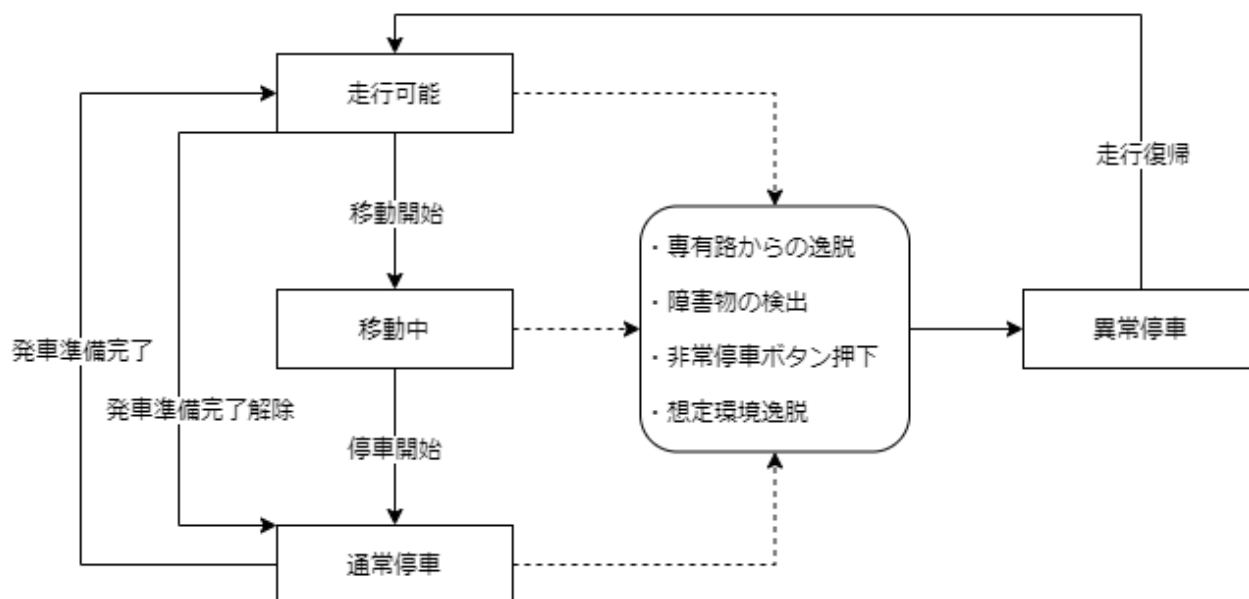
本システムは以下の登場人物によって構成される。

車両	自律的に車両の速度や角度を調整し、専有路上の白線に沿って走行する。本システムでは一台のみ運用する
乗客	車両に乗車している利用者
利用客	駅で車両を待っている利用者
駅	車両への乗降が行える場所。本システムでは駅A、駅B、駅Cの3カ所を設置する。ホームドアにより乗降時以外に利用客が専有路に出る事を防止する
インフラセンサ	複数台設置したカメラによって専有路の全域を監視し、管制に映像を送る
運用員	車両の走行に支障が出る障害物の除去や、車両が非常停車した場合の確認および後述する利用環境から逸脱していないかを監視する
管制	車両・インフラセンサ・駅・運用員の機能を統括してコントロールする
障害物	部外者や動物、その他の物体など専有路上に発生すると想定される走行の妨げになるもの
専有路	車両が走行する専有路

本システムは閉鎖された限定区間を想定するため一般車両や一般人（乗客・利用客や運用員を除いたシステム利用外の人を指す）はシステムの登場人物として記載していない。一般車両については進入も想定していないが、人や動物については侵入の想定はしている。一方で乗客・利用客・運用員はシステムとして保護する対象と定義しているが、侵入者については保護対象外と定義している。保護対象であるかどうかによって変わるのは、衝突が発生した場合にサービス事業者のようにシステムを運用する側が責任を持つかどうかであり、侵入者であっても衝突回避機能の対象となる。

b. システムの状態の整理

本システムは車両を主役としたシステムであり、車両の状態を基本としてホームドアの制御や車両への走行・停車指示などの制御内容を決定する。以下にシステムの状態と遷移を整理した。



通常は走行可能、移動中、通常停車のサイクルで運用され、非常事態（専有路からの逸脱・障害物の検出・非常停車ボタン押下・想定環境逸脱）が発生した場合には異常停車状態に遷移する。異常停車した場合には運用員が対応を行い、走行復帰をさせる事で走行可能状態になり通常のサイクルを再開する。この状態遷移を基にユースケースを策定する事でシステムの状態を網羅することができた。

4.1.1.2. 人工知能搭載 MaaS システムのユースケース整理

本システムを運用する中で発生する 13 通りのシチュエーションをユースケースとして下表に整理した。

車両の走行
車両の専有路逸脱停車
車両の衝突回避停車
車両の乗客操作での非常停車
車両の駅での通常停車
車両のその他の非常停車
駅での車両乗降
駅以外での車両乗降
車両の駅での通常発車
車両の走行復帰
障害物対応
非常停車対応
想定環境逸脱監視

表 ユースケース一覧

ユースケースの整理では車両の状態を基に各シチュエーションにおいて車両はどのような動きを

しているかという観点で考慮し、車両の走行や停車、停車時の車両乗降をユースケースとして挙げている。また、障害物の発生や専有路からの逸脱、非常停車操作などの駅で通常に行う停車とは異なる停車を行った場合に、運行を再開できるように対応するシチュエーションも挙げている。このユースケースについては後述する認証機関レビューにてシステムを網羅したユースケースの策定ができていたとの評価を得る事ができた。

4.1.1.3. ハザード分析手法

MaaS システムのハザード分析を STAMP (Systems-Theoretic Accident Model and Processes : システム理論に基づくアクシデントモデル) /STPA (System-Theoretic Process Analysis : システム理論に基づくプロセス分析) を用いて実施した。

本来の STAMP/STPA の手順ではアクシデント・ハザード・安全制約の識別」「制御構造モデルの構築」「非安全な制御の抽出」「ハザード要因の特定」の 4 つのステップで分析を行うが、本分析では独自の分析手順としてユースケースと組み合わせたハザードシナリオの導出を分析ステップに取り入れている。以下に本来の分析ステップと本研究との分析ステップ比較を示す。

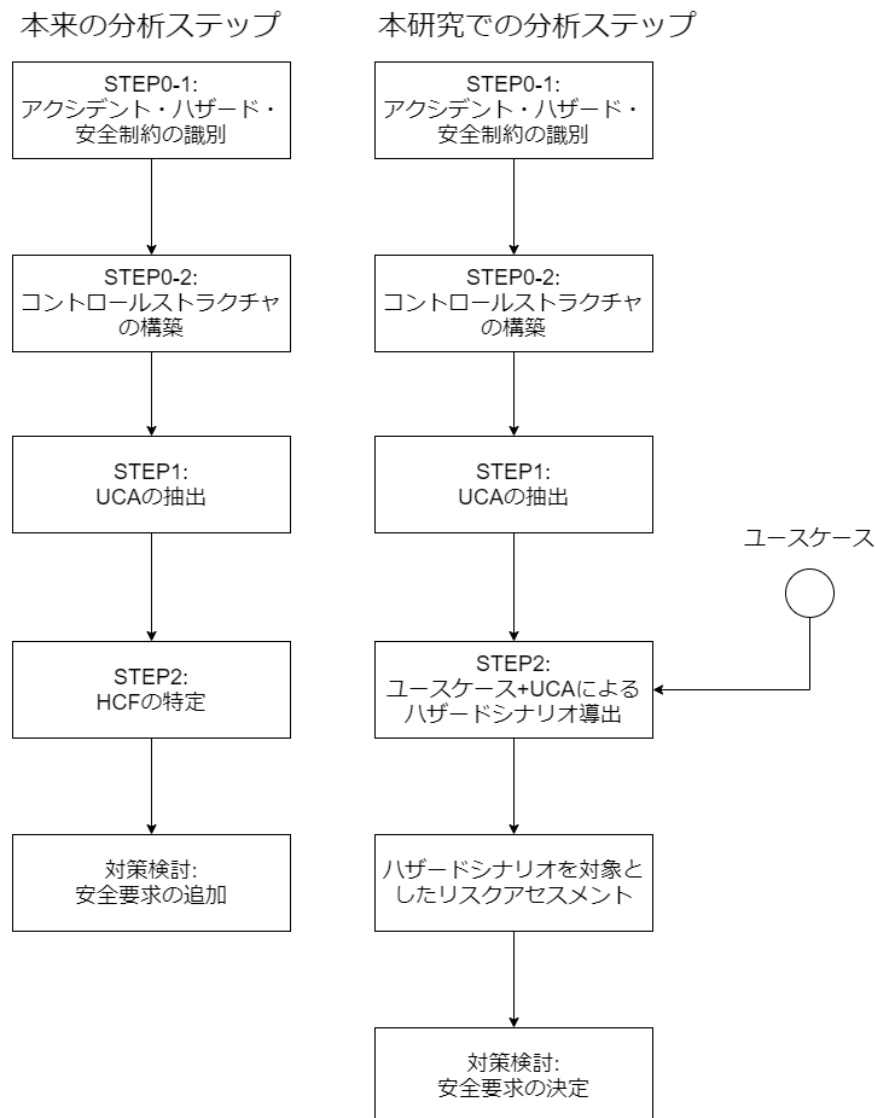


図 従来の分析ステップと安全コンセプト文書の分析ステップ比較

STAMP/STPA の本来の STEP2 で行うハザード原因 (HCF) の特定では、ガイドワードがアルゴリズム不備やアクチュエータ不備といった部品やプログラムの品質といった観点で設定されていて、事後解析的に分析をする場合には原因を明確化できるが、要件定義やシステムの定義段階では事前予測的な

分析しか行えず、原因特定より重要度が高いのはシチュエーションの決定である事が判明した。しかし、STAMP/STPA の本来の手順にはシチュエーションの決定は手順に入っておらず他の手法と組み合わせる必要がある。そこで本研究ではシチュエーションから策定したユースケースを骨組にし、その中で行われているコントロールアクションを考慮して UCA を当てはめていく分析手法を考案し実施した。

a. システムの機能を網羅されたコントロールストラクチャの構築

コンポーネント間の振る舞いは機能を行わせるコントロールアクションと、その結果を返すフィードバックで成り立つ。この組み合わせによってシステム全体の振る舞いを表現するものがコントロールストラクチャである。本年度の活動ではユースケースの検討を進める中でやり取りをするコンポーネントを洗い出し、コントロールストラクチャに反映させる手順を繰り返す事でコントロールストラクチャの精度を向上させ、網羅的なコントロールストラクチャの構築をすることができた。以下に構築したコントロールストラクチャを示す。

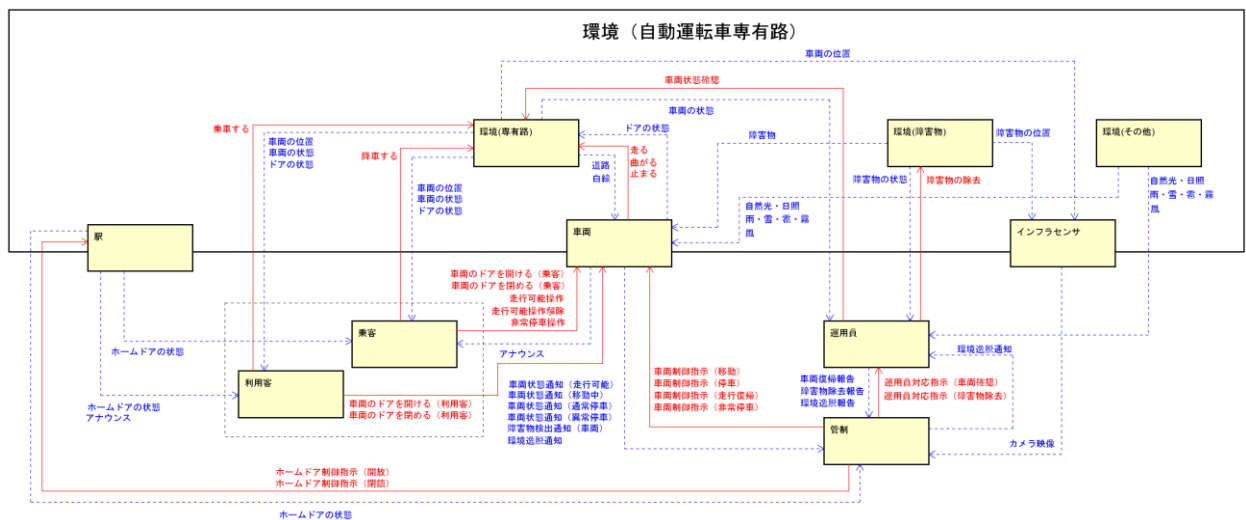


図 コントロールストラクチャ

b. ユースケースとコントロールループの関連付け

コントロールループとは STAMP/STPA の実施手順の中で抽出し、特定のコントロールアクションについてコントローラとコントロール対象に着目したコントロールストラクチャの一部である。コントロールループの分析によってコントロールアクションを行うトリガとなる入力や、コントロールアクションによって機能を実行した結果がどのコンポーネントに対して影響するのかを明確にすることができる。以下にコントロールループの例を示す。

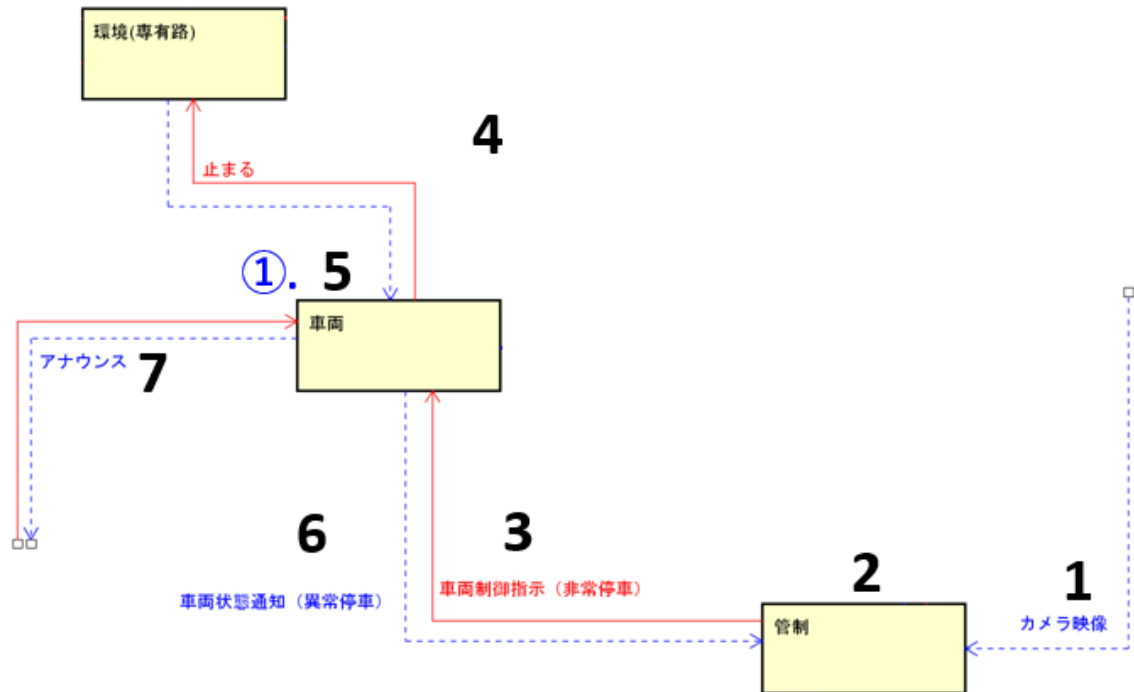


図 コントロールループの一例

この例では車両制御指示 (非常停車) がコントロールアクションでありその内容は車両に対して停車という動作を行わせる振る舞いである。この場合においてはコントローラが管制になり、コントロール対象が車両になる。図のみではシーケンスを把握しにくい為、図に採番を加える加工を行う事で可読性の向上を実現した。丸で囲った番号は1の開始前に満たしている必要がある事前条件を示す。以下にコントロールループの採番に従って動作を記述した例を示す。

① 車両状態が移動中となっている。

1. 入力としてカメラ映像が管制に対して送信される。
2. 管制は受信したカメラ映像に対して画像認識を行い、非常停車指示の要件を満たしているかを判定する。
3. 要件を満たした事をトリガとして管制から車両へ車両制御指示 (非常停車) を行う。
4. 指示を受けた車両は実際に停車を行う。これは、専有路上で車両自身を停車させた状態になるため、環境 (専有路) に対して行った「止まる」というコントロールアクションと解釈する。
5. 車両が車両制御指示 (非常停車) によって停車したことで車両状態 (4. 1. 1. 1. b 図 システムの状態遷移を参照) が移動中から異常停車へ遷移する。
6. 車両の状態が変化したため、管制に対して現在の状態を通知する車両状態通知 (異常停車) が行われる。
7. アナウンスによって車両内に状況のフィードバックを行う。

この例のようにコントロールループによってコンポーネント間で行われる振る舞いをシーケンスに起こす事ができ、カメラ映像に不備があったらどうなるか、アナウンスに不備があったらどうなるか、指示を受けた車両が止まらなかったらどうなるかといった影響分析の検討を進める事ができる。一方でシステム全体を示すコントロールストラクチャに対して同様の分析を行おうとした場合、複雑なシステムでは相関関係が入り組んでしまい効率的な分析が難しくなることが明確になった。その対策としてコントロールループによって一点に着目し、その中での影響を明確化した上でシステムの各コントロールアクションに対して併行して同様の分析を行った事でシステムを網羅した分析が実現できた。

4.1.1.4. ハザード分析の実施

a. アクシデントの基準の決定

安全コンセプト文書における安全とは ISO/IEC GUIDE 51:2014 に従い、「許容できないリスクがないこと」と定義している。この許容できるかどうかの基準を定める為にシステムにとって望ましくない事象を挙げ、リスクの等級分けを実施した。以下に本システムにおける定義を示す。

L1	生命の喪失または重傷者
L2	車両または車両外の物への損害・軽傷者
L3	時間的・経済的な損失

本研究のハザード分析においてはこの内 L1～L2 の基準を望ましくない事象（アクシデント）と定めた。サービスの品質の観点を考慮すれば L3 の時間的・経済的な損失も含める必要があり、L1～L2 をアクシデントとするのは本研究の安全コンセプト文書上での基準である。

b. アクシデント・ハザード・安全制約の決定

アクシデント基準の決定をしたら、その基準に基づき本システムにおいて発生しうるアクシデントを導出した。導出には登場人物一覧を基にして登場人物間で発生する事象にアクシデント基準に至るものがあるかを観点にして決定した。ハザードとはアクシデントが発生する状況を指し、安全制約はハザードを起こさない為に必要な状態を指す。以下にアクシデント・ハザード・安全制約の表を示す。

アクシデント	ハザード	安全制約 ID	安全制約
自動運転車両が利用客と衝突する	車両の走行中に専用路内に利用客がいる	SC1	車両の走行時に専用路内に利用客がいないこと
	車両が利用客の前で停車できない	SC2	車両が利用客の前で停車できること
自動運転車両が運用員と衝突する	車両の走行中に専用路内に運用員がいる	SC3	車両の走行中に専用路内に運用員はいないこと
	車両が運用員の前で停車できない	SC4	車両が運用員の前で停車できること
自動運転車両が専用路内に存在する障害物に衝突する	車両の走行中に専用路内に障害物がある	SC5	車両の走行中に専用路内に障害物が無いこと
	車両が専用路内に存在する障害物の前で停車できない	SC6	専用路内に存在する障害物の前で停車できること
自動運転車両が専用路外に存在する障害物に衝突する	車両が専用路から外れて走行する	SC7	車両は専用路から外れたまま走行しないこと
	車両が専用路外に存在する障害物の前で停車できない	SC8	専用路外に存在する障害物の前で停車できること
その他の環境条件逸脱による事故	その他の環境条件を逸脱している	SC9	想定する環境条件を逸脱した状態で車両が運行しないこと

表 アクシデント・ハザード・安全制約

アクシデント一つに対して二つのハザードを設定しているものは、発生と対策の不備という二段階で考慮している。システムを定義する中で発生が予想されているリスクについては対策もシステムの中で定義している。一方で、対策が不備によって機能しないケースは考慮する必要があり、ハザード

の条件として設定している。

c. UCA の抽出

アクシデント・ハザード・安全制約を決定したらコントロールストラクチャに示されている各コントロールアクションについて下記の4つのガイドワードに基づいて非安全になるかを分析する。

- ・与えられない (Not Providing)

そのコントロールアクションが行われない事によってハザードになるかという観点で分析する。

例えば車両が止まらない事は衝突の可能性があるためハザードとなるが、走らない事は衝突も起きない事になるので、ハザードとはならない。

- ・不正確に与えられる (Providing causes hazard)

そのコントロールアクションが本来行われないはずのタイミングで行う事によってハザードになるかという観点で分析する。

例えば駅のホームドアは車両が停車している時に開くのが本来のタイミングであり、車両がいない時に開いてしまうと車両が走っているにも関わらず人が専有路に出る事が出来てしまう為ハザードとなる。

- ・早過ぎる、遅延する (Too early / Too late)

タイミングを定められているコントロールアクションが本来のタイミングに対して早過ぎる、あるいは遅延したタイミングで実行するとハザードになるかという観点で分析する。

本来行われないタイミングは不正確に与えられる (Providing causes hazard) ケースで分析をする為、この観点における早過ぎる場合とは時限を設定しているものについてその時限到来前に実行される場合を指す。

例えば、車両が駅から発車する際にはホームドアを閉めるのにかかる時間を考慮して、10 秒待機してから発車するのが本来の時限である場合、5 秒しか経過していない時点で発車してしまうのは早過ぎる実行となる。この時、ホームドアがまだ閉まっている途中であり専有路に人が残っている可能性があるためハザードとなる。

- ・実行が短い、長い (Stop too soon / Applying too long)

コントロールアクションの実行に実行時間が定められているものについて本来の時間よりも短い、あるいは長い時間の実行によってハザードになるかという観点で分析する。

例えば、駅のホームドアに対して開く制御を行うと 30 秒の間開くのが本来の実行時間である場合に、20 秒しか開かずに閉じ始めるのは短い実行時間となる。この場合は閉じるのが早くても車両は本来開いている 30 秒の間は停車している前提であり、20 秒で閉じ始めても車両は停車のままなのでハザードとはならない。通信のように受け取ると実行するものについては実行時間の長短によってハザードとはならない。

このように、システムで定義した全てのコントロールアクションについて 4 種のガイドワードを適用し、安全制約に違反する場合は UCA、違反しない場合は非 UCA として仕分けする。以下に本システムにおける UCA の抽出例を示す。

No	CA	From	To	CA提供条件	Not Providing	Providing causes hazard	Too early / Too late	Stop too soon / Applying too long
1	走る	車両	環境(専有路)	[ユースケース_1,9,10] 1.車両状態が走行可能になっている 2.管制から車両制御指示(移動)を受ける	走らない	(UCA1-P-1) 車両状態が走行可能以外の状態で走る [SC9] (UCA1-P-2) 管制から車両制御指示(移動)を受けていない状態で走る [SC1][SC3][SC5][SC9]	遅延して走る	走行時間が長い
2	曲がる	車両	環境(専有路)	[ユースケース_1] 1.車両が走っている状態 2.車両LKAS機能によってカーブを検出する 上記が揃った場合	(UCA2-N-1) 曲がらない [SC7]	車両が走っていない状態で曲がる (UCA2-P-1) 車両LKASによってカーブを検出していない状態で曲がる [SC7]	(UCA2-T-1) カーブを曲がるのが早い [SC7] (UCA2-T-2) カーブを曲がるのが遅い [SC7]	(UCA2-D-1) カーブを曲がる時間が短い [SC7] (UCA2-D-2) カーブを曲がる時間が長い [SC7]

図 UCA 抽出の一例

d. ユースケースと UCA を組み合わせたハザードシナリオの検討

4.1.1.2 で示したユースケースでは、システムが正常に振る舞う事を想定している。そこに c. で抽出した UCA を組み合わせ、UCA の影響を考慮したシナリオをハザードシナリオとして導出することで非安全動作のパターンごとのシナリオの導出を実施した。以下にフローチャートを基にしたシナリオ分岐の例について示す。

ユースケース_1_フローチャート図

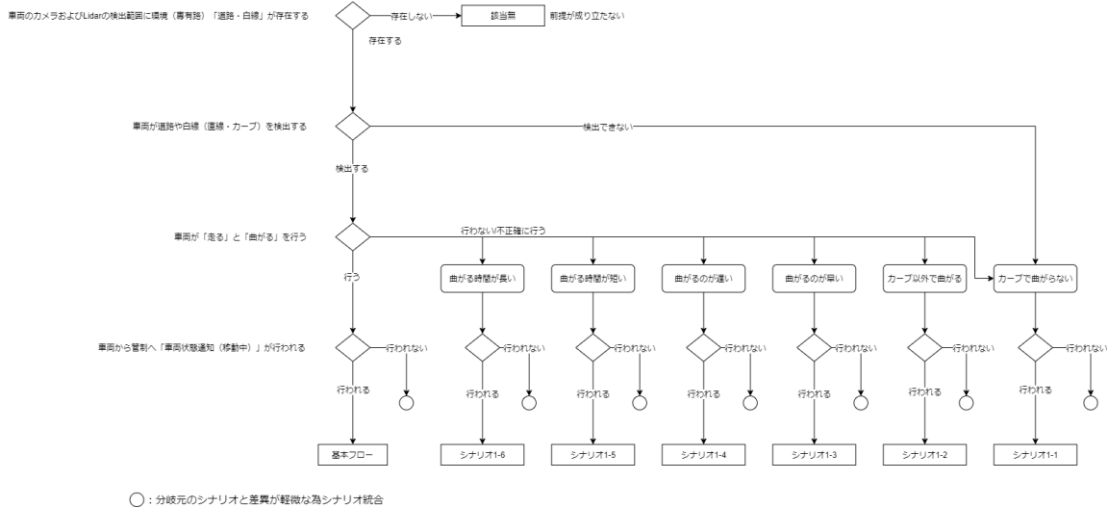


図 シナリオの分岐例

ハザードシナリオ単位をリスクの評価対象と定め、シナリオへの影響が軽微な分岐についてはシナリオを統合し、リスク評価の対象となるシナリオの基準を定めた。

4.1.1.5. ハザード分析の結果

4.1.1.4. ハザード分析の実施に示した手順によって 43 シナリオを導出した。

シナリオを整理する中で車両の止まる機能に関する不備が影響するシナリオが多く導出されるという傾向が見られ、対策を要する可能性が高いのがブレーキの振る舞いである事が明確にできた。一方で STAMP/STPA はシチュエーションが発生する可能性やその危険度、回避性については考慮しておらず、リスクの重み付けはリスク評価によって行う必要がある。

リスク評価を経て機能を追加する場合は再度ハザード分析の手順を実施し、影響を確認する。本研究ではシステムの中の特定のコンポーネント間で起きる振る舞いを見つける STAMP/STPA の強みとユースケースと組み合わせる事による網羅性を両立した分析手法の確立ができた。

4.1.1.6. 認証機関による技術レビュー

作成した安全コンセプト文書について、認証機関 SGS との技術ミーティングを開催し、レビューを受けた。

レビューの総評を以下に示す。

SGS 総評

- 出来ていること
 - 前提条件が明確になっており、網羅性の高いコントロールストラクチャをベースに、一貫性のある安全コンセプトとなっている。
- 改善点
 - 重要事項の漏れはないが、効果確認の手法や設計検証における優先度/基準を検討頂くとよい。

3

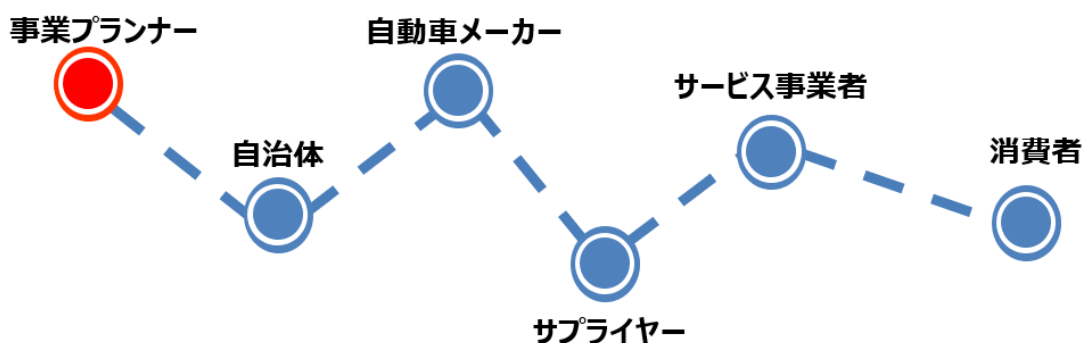
SGS の技術ミーティングでは、以下の結果となった。

- ・独自の手順を取り入れた STAMP/STPA の分析について方向性が正しいと認めてもらう事ができた。
- ・昨年度の課題であったシステム定義の明確化についても前提条件として明確化されている事を認めてもらうことができた。
- ・一方で、ハザードシナリオの選定基準や後続の分析としてハザード原因を特定した際の設計検証における優先度や基準について明確化されているとより品質が向上するとの意見をいただいた。

4.1.2. 人工知能搭載システムの安全性説明手法の確立

4.1.2.1. 安全性説明の必要性

人工知能搭載システムを事業化し、自動運転車両によるサービスを立ち上げるには自治体やメーカー、サプライヤーなど様々な団体（組織・個人を含む）が関わる事になる。以下に、事業プランナーの立場からサービスを提案する事を想定した場合に関わる団体の一例を示す。



一方で、関わる団体全てがシステムへの知識を有しているわけなく、専門用語を用いる安全コンセプトやその根拠である分析結果等を示しても理解を得られない可能性がある。

アクシデント・ハザード・安全制約の識別

アクシデントの内容によって統合できる箇所は統合を行う

アクシデント	ハザード	安全制約
A1 自動運転車が利用客と衝突する	車両の走行時に周囲に利用客がいる	車両の走行時に専有路上に利用客がないこと
A2 自動運転車が歩行者と衝突する	車両の走行時に歩行者がいる	
A3 自動運転車が専有路内に存在する障害物に衝突する	車両の走行時に障害物がある	
A4 自動運転車が専有路外に存在する障害物に衝突する	車両が専有路外に存在する	
A5 自動運転車が歩行者と衝突する	歩行者が専有路内に存在する	

コントロールストラクチャ

システム定義にて構築したコントロールストラクチャ

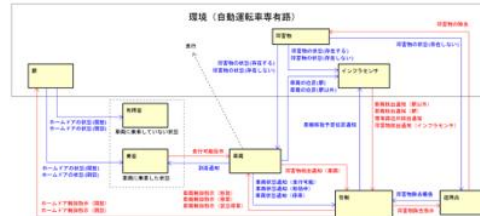


図 安全コンセプト文書や分析結果の一部

サービスを立ち上げ、運用していく上で安全性への理解は必要なものであり、システム知識を有さない相手に対しても伝わりやすい説明手法を確立する必要がある。そこで、本年度の活動では説明対象の明確化と図を用いたビジュアル化による説明手法の確立を行った。

4.1.2.2. 説明する対象と内容の分析

自動運転システムを事業化する際の関係者をステークホルダー（利害関係者という意味で企業・行政・NPO等の利害と行動に直接・間接的な利害関係を有する者を指す）として、それぞれに対して説明する対象とすらかどうかを分類した。以下にその表を示す。

ステークホルダー	例	説明の要・不要	理由
事業プランナー	(安全分析当事者)	－	今回の提案側の位置付けの為不要
自治体	〇〇市	要	自動運転を導入していただくために必要
サービス事業者	〇〇自動運転(株)	要	自動運転を導入していただくために必要
一般消費者	乗客	要	自動運転に理解をしていただくために必要
自動車メーカー	(株)〇〇自動車	不要	技術文書をそのまま理解できるため不要
サプライヤー	(株)〇〇部品製造	不要	技術文書をそのまま理解できるため不要

自動運転に関わるステークホルダーとしては自治体、サービス事業者、一般消費者の3者を説明対象者とした。しかし、自治体と一般消費者については「都市」「地方郊外」「観光地」によって動機が大きく異なり一括りに考慮できないため、ペルソナ（マーケティングにおいてイメージを共有する事を目的として設定する架空の人物像。顧客の象徴的なユーザー像などを具体化していく）を設定し、それぞれに対して導入の動機などを導いた。以下にその表を示す。

ステークホルダー	ペルソナ	導入の動機	導入の効果	注意すべき点	気にしていること（目論見）
自治体（都市）	大都市近郊型・地方都市型（国土交通省 先行モデル事業概要での分類）	免許を返納した高齢者の移動手段 渋滞解消	高齢者の免許返納による事故減 渋滞解消	訴訟	どういう危害が起こりうるか どういう対策を取っているか 定量的な話はしない

自治体 (地方 郊外)	地方郊外・過疎 地型	鉄道やバスが廃 止され、自家用 車を持たない住 民の移動手段	地方の活性化	訴 訟 経 済 面	どういう危害が起こりうるか どういう対策を取っている か 定量的な話はしない
自治体 (観光 地)	観光地型	観光客の移動手 段	観光客の増加	訴 訟	どういう危害が起こりうるか どういう対策を取っている か 定量的な話はしない
サービ ス事業 者	地方のバス会 社。創業 50 年。 運転手不足。 MaaS＝自動運転 くらいの意識。	運転手不足の解 消 コスト削減	運転手の労働環 境改善 利益率アップ	経 済 面	リスクと対策とそれによるコ スト感(定量的な話) 対策にどれくらいお金が掛 かるか 対策は事業プランナーが やること
一般消 費者 (都市)	都市の高齢者。 車の免許は返納 済み	車の免許返納 買い物の交通手 段 高齢者の外出支 援	買い物が容易に なる 高齢者が健康に なり、健康保険費 抑制	受 傷	事故が起きないこと 「安全です」というアピール 定量的な話、具体的に何を やっているかまでは話さな い
一般消 費者 (地方 郊外)	地方の過疎地の 高齢者。車の免 許は返納済み 地方の過疎地の 学生。交通手段 がない	車の免許返納 買い物の交通手 段 通学の交通手段	買い物が容易に なる 通学が容易にな る	受 傷	事故が起きないこと 「安全です」というアピール 定量的な話、具体的に何を やっているかまでは話さな い
一般消 費者 (観光 地)	観光客	知らない土地でも 目的地に行ける ようになる	地元詳しくない	受 傷	事故が起きないこと 「安全です」というアピール 定量的な話、具体的に何を やっているかまでは話さな い

4.1.2.3. 表現方法の分類

表現の方法については構造や状況等を俯瞰して描く俯瞰図、シチュエーションをイメージする為のイラスト、動的かつ詳細な説明をすることができる動画による説明が考えられる。

以下にそれぞれの方法の特徴を示す。

a. 俯瞰図

システム概要のイメージ化など、広い視野で見るものの図式化を行う手法。全体を見通す必要があるものの説明に用いる。

閉鎖区間における無人「人・モノ」輸送サービス

無人車両による限定区間内ピストン輸送システム

管制が各機能と通信を行い運行を管理する事で専用区間内における自動運転車両の
自律的自動運転を支援し、安全な輸送を行うことを目的としたシステム

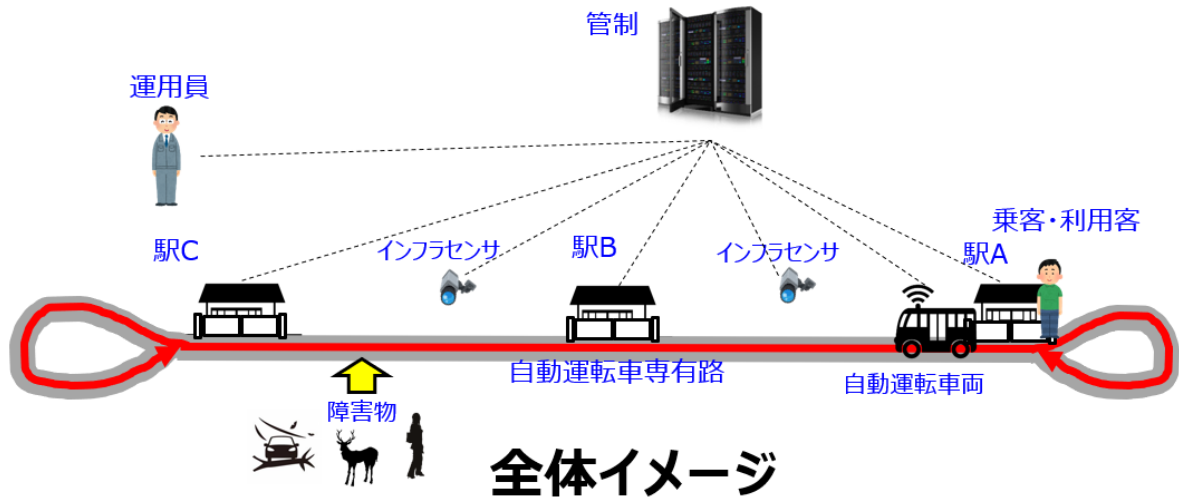


図 4.1.1.1 人工知能搭載 MaaS システムの定義で示したシステム概要図

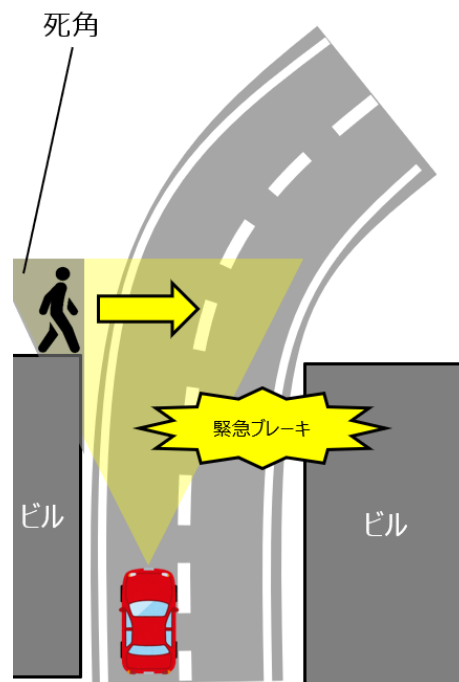


図 事故シチュエーションの俯瞰図

俯瞰図を用いるメリット・デメリットを以下に示す。

- ・メリット

概要を一枚絵で伝えられるため文章だけでは理解しにくい内容を伝えやすい
アニメーションを加える事である程度動的な状況説明に用いる事もできる

- ・デメリット

図の複雑さによっては内容を読み取るのに時間を要する場合がある
詳細な情報を含めると視認性が損なわれる

b. イラスト

特定のシチュエーションや意味を図式化する手法。標識や注意喚起などに用いる。



図 標識の一例



図 注意喚起イラストの例 1



図 注意喚起イラストの例 2

俯瞰図とは逆の特徴があり、情報をシンプルにすることで特定の事象を伝えるのに向いている手法。

用いるイラストも視認性が高く瞬間的に読み取れるものが多い。一方で動的な説明には用いりにくい。イラストを用いるメリット・デメリットを以下に示す。

- ・メリット

文字をほとんど用いらずに直感的に内容の理解が可能。

伝わりやすいデザインの考案は検討が必要だが、作成自体は容易であり専門的なスキルが不要。

- ・デメリット

一つのイラストにつき一つの内容しか含まないものが多く、単一イラスト毎の情報量が少ない

上記の為に複数の事項を伝える場合にはその数量分の枚数が必要になり、類似していると混同のリスクが発生する。

文字を用いない標識のようなものの場合、その意味を見る側が記憶している必要があり意図通りに伝わらない可能性がある。

c. 動画

時間軸を含めた動的な状況説明や、シミュレータ等を用いる事によるより現実に近いイメージでの表現、俯瞰図やイラストでは表現しきれない情報量の多い事象を図式化する手法。



図 3D モデルを用いた事故の再現動画のイメージ

(引用元: 3D 交通事故レポートメーカー <https://www.megasoft.co.jp/accident/>)



図 シミュレータを用いた動画のイメージ

複雑な状況の説明や動的に見せる事でイメージをしやすくなるような事象の説明に用いやすい。特に自動運転車両の場合、公道での走行には法律による大きな制限があり実験的な走行が難しい事や、管轄団体の許可申請が通りにくい等の事情があり、実車による撮影を行いにくい。また、システムの検証等安全性の保証ができていない段階で実車の走行を行う事はリスクが高く、提案段階で実車を用いた資料映像が準備できない場合に仮想環境シミュレータで代替する事でイメージを説明するという方法が可能になる。

動画を用いるメリット・デメリットを以下に示す。

- ・メリット

動的で複雑なシチュエーションを発生、再現させることができる。

自動運転シミュレータを一度用意する事でシチュエーションに対応した資料を用意しやすくなる
(俯瞰図やイラストならその都度新たに用意する必要がある)

動画内に説明を付与させる事で見る側の知識に頼らずに伝える事ができる

- ・デメリット

30 秒の動画なら 30 秒視聴しないと内容を把握できない為、情報取得にかかる時間が長い。

上記の特徴から通りがかりに見せるといった状況に用いると内容が伝わらない可能性がある。

作成には専門知識が必要であり、工数もかかる方法である。

4.1.2.4. 自動運転シミュレータ

4.1.2.3.c の動画の内、自動運転シミュレータは複雑な自動運転システムの説明に有用な方法であると考えられ、具体的に用いる事ができるシミュレータを検討した結果 AirSim によるシミュレーションが有用である事が明確化した。AirSim は Microsoft AI & Research Group が開発・配布しているオープンソースシミュレータ。当初はドローンのシミュレーションのための物であり名称の AirSim は空中を飛ぶドローンに由来している。シミュレート対象をドローンから他の乗り物に変える事も可能であり自動車もシミュレート対象になっている。

AirSim の特徴を以下に示す。

- ・プラットフォームは Windows 用、Linux への対応も予定している。
- ・ゲームエンジンを使用し、Unreal Engine 版と Unity 版がある
- ・本体は MIT ライセンス (ゲームエンジンのライセンスは別途)
- ・C++ または Python で自動運転プログラムが書ける。100 行程度のプログラムで走らせることができる (サンプルあり) → プログラミングの粒度は「ハンドブレーキは false、スロットルは -0.5、ステアリングは 1」程度
- ・ネットワーク経由での制御も可能
- ・複数の車にも対応。今回のシミュレーションに適している
- ・マニュアル操作も可能 (キーボード/ジョイスティック/ドライブ用コントローラなど)

- ・アセット（資産。車両やコースの 3D データ）が豊富
- ・Unreal Engine 版が本家→Unity 版に移植中（現在β版）

本年度の活動で AirSim を動作させた環境を以下に示す。

- ・ノート PC (CPU : Core i5/2.8GHz、GPU : オンボード)

このスペックの場合には毎秒 5 枚程度の品質となる。画質を落とすことで速度を確保することは可能であり、毎秒 20 枚～30 枚程度の品質を実現した。

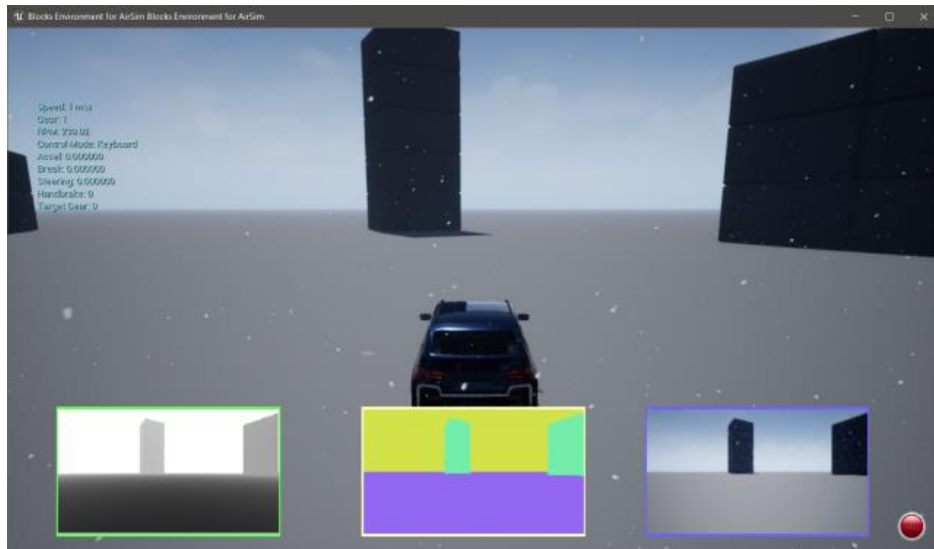


図 AirSim の動作イメージ

```
std::cout << "Press Enter to take turn and drive  
backward" << std::endl; std::cin.get();  
controls.handbrake = false;  
controls.throttle = -0.5;  
controls.steering = 1;  
controls.is_manual_gear = true;  
controls.manual_gear = -1;  
client.setCarControls(controls);
```

図 AirSim を動作させるコードの一例

コーディングに用いる C++ または Python の言語知識に加え、Unreal Engine の仕様を理解している必要があり導入敷居は低くないものの、導入の効果は大きいことが明確になった。
主流は Unreal Engine 版であり、Unity 版は移植テスト版の為更新がされていない部分やアセットが Unreal Engine 版ほど充実していない事も明確になった。

4.1.2.5. 説明対象と表現手法

各表現手法の特徴から、時間をかけた説明が可能であるステークホルダーに対しては俯瞰図や動画を用いた説明、時間をかけた説明が難しい場合が考えられるステークホルダーに対してはイラストを用いるのが有用であると明確になった。

一例として、自治体やサービス事業者に対しては俯瞰図を含めた説明資料を用意し提案を行い、一般消費者に対しては注意喚起のステッカー等でシステムが安全に運用できるように逸脱してほしくない事象についてのルール周知を行う事で許容できないリスクの発生を低減させ、システムの安全性説明手法を確立できる事を提言できた。

自治体向け

想定される危険事象

- 車両との接触による事故
 - 乗客の負傷
- 車両が占有路外への逸脱
 - 乗客の負傷
 - 第三者の負傷

これらの危険事象に対して対応しなければならない

危険事象への対応策

- 事故を未然に防ぐために、本自動運転では複数の安全対策をとっています。
- 車両が障害物を発見した場合、その場で安全に停車します
- 市内のカメラが障害物を発見した場合、その場で安全に停車します
- 車両が道路から逸脱した場合、その場で安全に停車します

どういうハザードが起こりうるか

どういう対策をとっているか

定量的な話はしない

サービス事業者向け

想定されるリスクと対策

自動運転で想定されるリスク対策は以下の通りである

- 自動運転サービスを提供することで考えられるリスクとして、事故発生による経済的リスクが考えられる
- 事故発生を未然に防ぐために、車両側とインフラ側にセンサを設置する

車両側センサ

車両の自動運転、車載センサが必要

インフラ側センサ

道路を認識できるが、インフラ側センサが必要、運行直前・センサの検出範囲

必要とされるコスト

リスクに対応するためのコストは以下の通りである。

車両側センサの設置

- 車両側に障害物を検出するためのセンサを設置する
- 車両側センサ整備・台数
- ソフトウェア開発費用

インフラ側センサの設置

- インフラ側に障害物の侵入または車両の逸脱を検出、センサを設置する
- インフラ側センサ・台数
- ソフトウェア開発費用 (センサ側、制御側)
- 障害物検出用インフラ側センサと、車両逸脱検出用インフラ側センサは共用できる

リスクと対策

対策に必要なコスト感

具体的にどれくらいコストが必要か



4.1.3. 令和元年度の活動まとめ

本年度の研究活動によって人工知能搭載システムに対しても適用が可能な安全分析手法を提言することができた。

また、以下の取り組みについて本研究の強みである事が認識できた。

<本研究の強み サブテーマ 1>

STAMP/STPA を用いた分析の具体的な検討例
ユースケースとハザード分析結果を組み合わせた網羅的検討方法
自動運転レベル 4 を想定した MaaS システムに対する分析
複雑なシステムを説明する為のビジュアル化による説明手法

<改善点>

ミスユースや性能限界を検討していくフェーズでのハザード原因特定手法の検討
コントロールストラクチャより可読性の高い機能とエレメントの配置図の構成
ビジュアル化による具体的な提案円滑化の事例

4.2. 人工知能搭載システムの潜在リスクへの対応（サブテーマ 2）

令和元年度のサブテーマ 2 の活動では、自動運転レベル 4 システムの安全コンセプトを再構築し、認証機関による内容妥当性確認を受け、本活動の方向性の正しさを確認した。

また、それらの知見を人工知能搭載システム安全ガイドライン「SEAMS ガイドライン」に反映し、第 2 版を完成した。

作成した SEAMS ガイドラインの目次を以下に示す。

1. 本ガイドラインについて
 - 1.1. 目的
 - 1.2. 対象範囲
 - 1.3. 想定読者
 - 1.4. 本ガイドラインによって改善できる AI の課題
 - 1.5. 本ガイドラインの位置付け
 - 1.6. QA4AI との関係
 - 1.7. 用語集
2. 背景
 - 2.1. 普及する AI 搭載システム
 - 2.2. AI の分類整理
 - 2.3. AI を搭載した安全関連システムの課題認識
 - 2.3.1. AI の信頼性
 - 2.3.2. 安全に影響のある AI の特性
 - 2.3.3. AI の構成要素と信頼性
 - 2.3.4. AI は機能安全規格では非推奨
 - 2.3.5. AI の安全立証における課題
 - 2.3.6. 説明可能性の高い AI
3. AI 搭載システムの安全設計パターン
 - 3.1. AI 搭載システムの安全設計における注意点や課題
 - 3.2. 方針 1: AI 自体を安全設計する
 - 3.2.1. 1a) 機能安全開発パターン
 - 3.2.2. 1b) 安全性評価パターン
 - 3.2.3. 1c) Proven-in-use パターン
 - 3.2.4. 1d) 多重化設計パターン
 - 3.2.5. ソフトウェアの不確かさへの対応
 - 3.3. 方針 2: AI は安全設計せず、外部に安全メカニズムを設ける
 - 3.3.1. 2a) 機能安全機能パターン
 - 3.3.2. 2b) 比較パターン
 - 3.3.3. 2c) 防御設計パターン
 - 3.4. 一時故障の影響と対策
4. AI 搭載システムを用いたサービスの安全開発プロセス
 - 4.1. ①サービスレベルのシステム定義
 - 4.2. ②ハザード分析及びリスクアセスメント
 - 4.2.1. アクシデント・ハザード・安全制約の決定
 - 4.2.2. コントロールストラクチャの作成
 - 4.2.3. UCA の抽出
 - 4.2.4. UCA を含めたユースケースでの非安全シナリオ導出
 - 4.2.5. リスクアセスメント
 - 4.2.6. 参考：シミュレーション環境による自動化
 - 4.3. ③サービスレベルの安全要求導出
 - 4.4. ④サービスレベルの安全要求割り付け
 - 4.5. ⑤システム単体のアーキテクチャ設計
 - 4.6. ⑥システムレベルの安全要求導出
 - 4.7. ⑦システム安全分析
 - 4.8. ⑧初期 AI コンポーネント開発
 - 4.9. ⑨AI 学習フェーズ
 - 4.9.1. Automotive SPICE のテラリング
 - 4.9.2. SOTIF における AI 学習ワークフロー
 - 4.9.3. 説明可能性の高い AI モデルワークフロー
 - 4.10. ⑩ハードウェア・ソフトウェア開発
 - 4.11. ⑪システムレベル統合試験
 - 4.12. まとめ

5. 付録 A: AI 関連資料の調査
 - 5.1. 既存のガイドラインの調査
 - 5.2. 先行研究の調査
 - 5.3. AI の研究開発の原則の策定
 - 5.4. QA4AI ガイドラインの概要
6. 付録 B: IEC 62998 の概要と AI 搭載システムへの適合
 - 6.1. IEC 62998 の概要
 - 6.2. IEC 62998 の特徴
 - 6.3. IEC 62998 における開発の流れ
 - 6.4. IEC 62998 の AI 搭載システムへの応用
 - 6.5. IEC 62998 の課題
 - 6.6. 参考情報: IEC 62998 を満たしているセンサの例
7. 付録 C: 説明可能性の高い AI とは
 - 7.1. 説明可能な AI (XAI)
 - 7.2. 説明可能性の高い AI モデル開発ワークフロー
 - 7.2.1. ワークフローの概要
 - 7.2.2. ①探索的データ分析 (EDA) & 視覚化
 - 7.2.3. ②ベンチマークや再現性の確立
 - 7.2.4. ③手動な、個人的な、まばらな、わかりやすい特徴選択
 - 7.2.5. ④公平性、プライバシー、セキュリティのための前処理
 - 7.2.6. ⑤制約された、公平な、解釈可能な、個人的な、シンプルなモデル
 - 7.2.7. ⑥予測の調整
 - 7.2.8. ⑦伝統的なモデル評価&診断
 - 7.2.9. ⑧説明
 - 7.2.10. ⑨モデルデバッグ
 - 7.2.11. ⑩社会的偏見の考査と修復
 - 7.2.12. ⑪リスクの定量化と計画
 - 7.2.13. ⑫人間のレビューと文書化
 - 7.2.14. ⑬配備、管理、監視
 - 7.2.15. ⑭自動化されたモデル決定のヒューマンアピール
 - 7.2.16. ⑮根本原因分析
 - 7.2.17. ⑯反復
 - 7.2.18. ⑰廃止
 - 7.3. QA4AI ガイドラインの活用
8. 付録 D: 関連標準の概要
 - 8.1. ISO/PAS 21448 (SOTIF)
 - 8.1.1. ISO/PAS 21448 のプロセス概要
 - 8.1.2. ISO 26262 との関係
 - 8.1.3. 自動運転レベル 3 以上への適用
 - 8.2. IEC 62853 (DEOS)
 - 8.3. IEC 60721
 - 8.4. 国や地域による安全の考え方の違い
9. 付録 E: 自動運転と安全
 - 9.1. SAE J3016 で定義されている自動運転のレベル
 - 9.2. 自動運転における「安全状態」の例
 - 9.3. 自動運転における「安全時間」の例
 - 9.4. 自動運転システムの構成要素
 - 9.5. 自動運転における一般的なセンサ・フュージョン
 - 9.6. 自動運転システムにおける環境条件の例
 - 9.7. 自動運転技術のリスクアセスメントの必要性
10. 付録 F: 具体的な適用事例

本サブテーマでは、以下の活動を遂行し、その知見を SEAMS ガイドラインとしてまとめた。

- ・安全認証に必要な各種ドキュメントのテクニカルレポート取得（サブテーマ 2-2）
 - ・具体的な自動運転レベル 4 システムのリスクアセスメントの実施
 - ・具体的な自動運転レベル 4 システムの安全対策の検討
 - ・具体的な自動運転レベル 4 システムの AEBS 機能に対する安全コンセプト文書の作成
 - ・認証機関による技術レビューならびに評価

- ・人工知能の種別・安全度・安全対策方法の分類（サブテーマ 2-3）
 - ・システム定義やハザード分析&リスクアセスメントの実施方法や観点の整理
 - ・人工知能搭載システムの安全論証手法の分類と具体的な手順の整理
 - ・人工知能搭載システムの安全説明力の高い開発プロセスの整理
 - ・人工知能搭載システム安全ガイドラインの対外紹介の実施

4.2.1. 安全認証に必要な各種ドキュメントのテクニカルレポート取得（サブテーマ 2-2）

4.2.1.1. 具体的な自動運転レベル 4 システムのリスクアセスメントの実施

サブテーマ 1 における MaaS 安全コンセプト「非安全動作（UCA）を含めたユースケースでのシナリオ導出」の結果を受け、MaaS システムに対するリスクアセスメントを実施した。

リスクアセスメントの手法として、ISO 26262-3 のパラメータであるシビアリティ (S)、曝露の確率 (E)、コントローラビリティ (C) 及び安全度水準 (ASIL) を用いた。具体的な評価基準を下表に示す。

リスクパラメータ		ISO26262 における記載	評価基準
曝露の確率 (E)	E0	極めて起こりにくい	・車両走行中の、乗客による故意の開錠、ドア開放
	E1	非常に低い確率 大多数のドライバにとっては、年に 1 回未満しか発生しない状況	・走行可能操作時のアナウンス後に利用客が専有路上にいる状態 ・駅正面の専有路上に車両がない時にホームドアが開放された場合の利用客の立ち入り ・車両制御指示（走行復帰）時のアナウンス後に運用員が車両前方にいる状態 ・意図しないタイミングでの運用員による障害物除去及び車両状態確認
	E2	低い確率 平均動作時間の 1%未満 大多数のドライバにとっては、年に数回しか発生しない状況	・専有路上の障害物の存在
	E3	中程度の確率 平均動作時間の 1~10% 平均的なドライバにとっては、月に 1 回以上発生する状況	—
	E4	高い確率 平均動作時間の 10%超 平均して、ほとんど全ての運転中に発生する状況	・単一のコンポーネント故障によりの発生する事象
コントローラビリティ (C)	C0	通常のコントロール可能	・C1 の回避手段のうち 2 つ以上に該当 ・非常停車操作による停車 ・シーケンス上、車両が発進しない状態
	C1	簡単にコントロール可能 平均的なドライバ又は他の交通当事者の 99%超が、危害を回避できる	・車両の障害物検出機能による衝突回避 ・車両の専有路逸脱検出機能による衝突回避 ・運用員または利用客の専有路立ち入

（令和元年度戦略的基盤技術高度化支援事業）
「自律的自動運転の実現を支える人工知能搭載システムの安全性立証技術の研究開発」研究成果報告

			り時の周囲確認
	C2	普通にコントロール可能 平均的なドライバ又は他の交通当事者の90%から99%が、危害を回避できる	—
	C3	コントロールが困難又コントロール不能 平均的なドライバ又は他の交通当事者の90%未満が、危害を回避できる、又は、なんとか回避できる	・車両が停車状態または異常停車状態での、意図しないタイミングの走行開始 ・非常停車操作または車両制御指示の不備による走行継続・駅の通過
シビアリティ (S)	S0	傷害なし	・20km/h 未満で走行時の利用客、運用員または障害物との衝突による、乗客への影響(車両の発進時、減速中) ・非常停車操作または車両制御指示の不備による走行継続・駅の通過 ・シーケンス上、車両が発進しない状態
	S1	軽度及び中程度の傷害	・10km/h 未満で走行時の利用客または運用員との衝突(車両の発進時)
	S2	重度及び生命を脅かす傷害(生存可能性あり)	・20km/h 未満で走行時の利用客または運用員との衝突(車両の減速中)
	S3	生命を脅かす傷害(生存不確実)、致命傷	・20km/h 以上で走行時の利用客または運用員との衝突(通常走行中) ・20km/h 以上で走行時の障害物との衝突による乗客への影響 ・20km/h 以上で走行時の乗客の車両からの転落

リスクアセスメントの結果、ASIL の付与されたシナリオを以下に示す。

シナリオ	E 評価		C 評価		S 評価		ASIL
運用員の対応中に車両が走行を再開する	E4	単一のコンポーネントの故障により発生する	C3	意図しないタイミングでの発車となるため、専有路上の運用員が回避できる可能性は低い	S1	車両の発進時のため、衝突した場合の衝突時の速度は10km/h 未満であり、運用員の傷害は軽傷であると考えられる	B
車両が障害物が発生しても止まらず衝突する	E2	毎日運用前に点検を行うため、専有路上に障害物が存在する可能性は低い	C3	障害物検出機能が故障した場合、車両及び乗客による回避は不可能	S3	車両が最高時速50kmで障害物に衝突した場合、乗客が死傷する恐れがある	B
車両が障害物を検出してから止まるまでが遅れ、衝突する	E2	毎日運用前に点検を行うため、専有路上に障害物が存在する可能性は低い	C3	障害物検出機能が故障した場合、車両及び乗客による回避は不可能	S3	車両が最高時速50kmで障害物に衝突した場合、乗客が死傷する恐れがある	B
車両が障害物を検出して止まるのに時間がかかる	E2	毎日運用前に点検を行うため、専有路上に障害物が存在する可能性	C3	障害物検出機能が故障した場合、車両及び乗客による回避は不可	S3	車両が最高時速50kmで障害物に衝突した場合、乗客が死傷する恐れ	B

		は低い		能		れがある	
乗降中に車両が発車する	E4	単一のコンポーネントの故障により発生する	C3	意図しないタイミングでの発車となるため、専有路上の利用客が回避できる可能性は低い	S1	車両の発進時のため、衝突した場合の衝突時の速度は 10km/h 未満であり、利用客の傷害は軽傷であると考えられる	B

4.2.1.2. 具体的な自動運転レベル 4 システムの安全対策の検討

前項にて ASIL の付与されたシナリオを持つ各 UCA について、リスクを回避するための安全策を、下表のように安全要求として定義した。その際、安全要求を車両のみに集約した。

UCA ID	シナリオ	ASIL	安全要求	安全要求 ID
UCA01	運用員の対応中に車両が走行を再開する	B	車両が停車状態の場合は、車両制御指示(移動)を受けても停止状態を維持し、走行を開始しない	MAAS-SR-01
UCA24	乗降中に車両が発車する	B		
UCA09	車両が障害物が発生しても止まらず衝突する	B	前方 30m に高さ 10cm 以上の障害物、人間、動物を検出した場合、3.5 秒以内に緊急停車する	MAAS-SR-02
UCA10	車両が障害物を検出してから止まるまでが遅れ、衝突する	B		
UCA11	車両が障害物を検出して止まるのに時間がかかる	B		

4.2.1.3. 具体的な自動運転レベル 4 システムの AEBS 機能に対する安全コンセプト文書の作成

サブテーマ 1 で定義した MaaS システム内の自動運転車両に搭載される AEBS 機能の安全コンセプトである、「AEBS 安全コンセプト」を作成した。
作成した安全コンセプトの目次を以下に示す。

1	概要
1.1	目的
1.2	用語定義
1.3	準拠する規格
1.4	関連文書
2	自動運転車両の機能
2.1	各機能の概要
2.1.1	AEBS
2.1.2	ACC
2.1.3	LKAS
2.2	利用環境
2.3	システム構成
2.4	車両レベルでの共通原因故障
2.5	自動運転車両に割り当てる安全要求
3	AEBS 安全要求
3.1	安全目標の詳細化
3.2	機能安全要求
3.3	安全分析結果
4	システムマティック故障対策
4.1	開発フェーズ全体像
4.2	①初期 AI コンポーネント開発
4.3	②コンセプトフェーズ
4.4	③システム設計
4.5	④ハードウェア・ソフトウェア開発
4.6	⑤AI 学習フェーズ
4.7	⑥統合試験

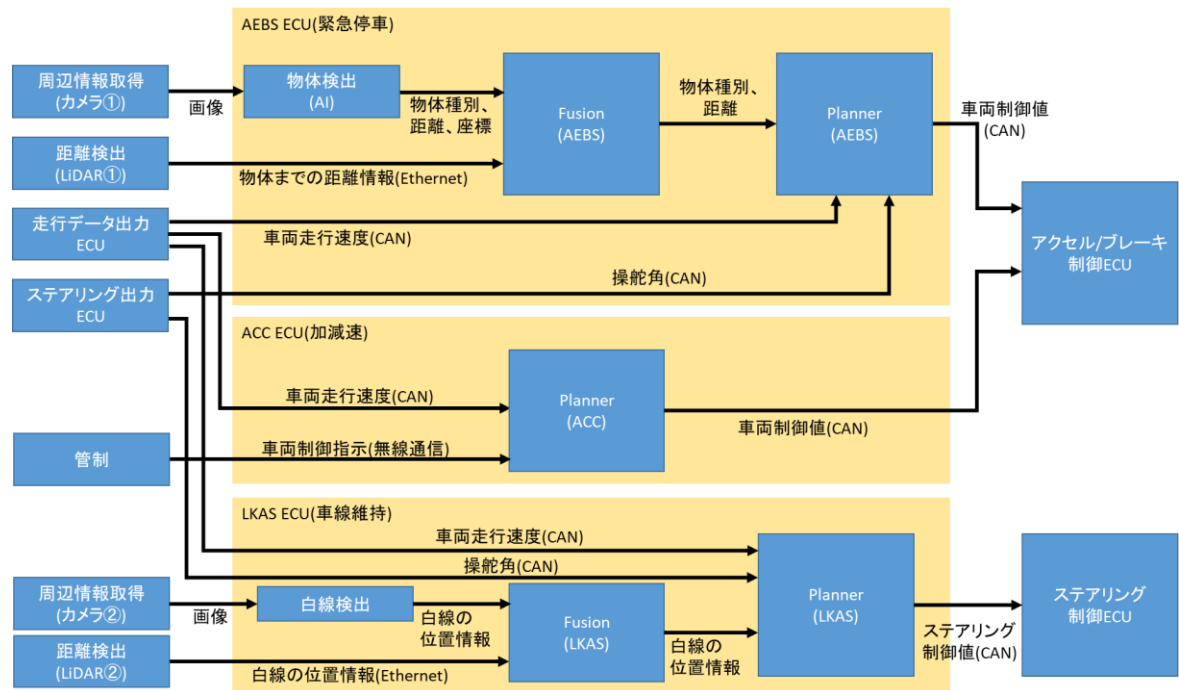
昨年度作成した ACC 自動運転システム安全コンセプトは高速道路環境を想定していたが、本年度定義した MaaS システム環境には車両が 1 台しか存在せず、車間維持を行う必要がない。そのため ACC システムには AI を搭載せず車両の発車、停車を司る機能と位置づけ、AEBS システムに AI を搭載することとした。

自動運転車両は以下の機能を持つ。

- AEBS：カメラ及び AI による物体検出と、LiDAR による距離検出によって専有路上の障害物を検出し、必要な際に緊急停車する機能
- ACC：管制からの車両制御指示に従い、車両の発車、停車を行う機能
- LKAS：カメラ及び LiDAR によって白線を検出し、ステアリング制御を行うことによって車両が白線から逸脱せず走行する機能

自動運転車両に搭載される AEBS、ACC、LKAS 機能のシステム構成を以下に示す。

(令和元年度戦略的基盤技術高度化支援事業)
「自律的自動運転の実現を支える人工知能搭載システムの安全性立証技術の研究開発」研究成果報告



上記のシステム構成で発生し得る共通原因故障について分析を行い、対策を安全要求として定義した。これと MaaS 安全コンセプトで定義した安全要求のうち、AEBS に割り当たる安全要求を、AEBS 安全コンセプトにおける安全目標として定義した。

安全要求 ID	安全要求	ASIL	割当先	安全目標 ID
MAAS-SR-01	車両が停車状態の場合は、車両制御指示(移動)を受けても停止状態を維持し、走行を開始しない	B	ACC	—
MAAS-SR-02	前方 30m に高さ 10cm 以上の障害物、人間、動物を検出した場合、3.5 秒以内に緊急停車する	B	AEBS	SG1
CAR-SR-01	走行データ出力 ECU の故障を検出した場合は非常停車する	B※	AEBS ACC LKAS	SG2
CAR-SR-02	ステアリング出力 ECU の故障を検出した場合は非常停車する	B※	AEBS LKAS	SG3
CAR-SR-03	アクセル/ブレーキ制御 ECU の故障を検出した場合は非常停車する	B※	AEBS ACC	SG4

安全目標SG1～SG4を詳細化し、システム構成で示した各機能ブロックに機能安全要求として割り当てた。

機能ブロック	機能安全要求 ID	機能安全要求	ASIL
周辺情報取得(カメラ①)	—	—	QM
物体検出(AI)	—	—	QM
距離検出(LiDAR①)	AEBS-FSR-DIS-001	前方 30m にある高さ 10cm 以上の物体を検出すること。	B
	AEBS-FSR-DIS-002	LiDAR で算出する距離情報の信頼性を担保	B

（令和元年度戦略的基盤技術高度化支援事業）
「自律的自動運転の実現を支える人工知能搭載システムの安全性立証技術の研究開発」研究成果報告

		すること。	
	AEBS-FSR-DIS-003	LiDAR の自己診断を行うこと。	B
	AEBS-FSR-DIS-004	LiDAR の自己診断結果を送信すること。	B
Fusion (AEBS)	AEBS-FSR-FUS-001	LiDAR の距離情報と AI の距離情報が不一致の場合、LiDAR の情報を優先すること。	B
	AEBS-FSR-FUS-002	障害物が検出範囲に入ってから 300ms 以内に障害物検出を行うこと。	B
	AEBS-FSR-FUS-003	Fusion で算出する距離情報の信頼性を担保すること。	B
	AEBS-FSR-FUS-004	LiDAR の自己診断結果を受信すること。	B
	AEBS-FSR-FUS-005	Fusion の自己診断を行うこと。	B
	AEBS-FSR-FUS-006	Fusion の自己診断結果を送信すること。	B
	AEBS-FSR-FUS-007	AI からの出力がない場合、Planner への出力を停止すること。	B
	AEBS-FSR-FUS-008	LiDAR が故障している場合、Planner への出力を停止すること。	B
	AEBS-FSR-FUS-009	LiDAR からの出力がない場合、Planner への出力を停止すること。	B
走行データ 出力 ECU	AEBS-FSR-FEB-001	走行データ出力 ECU で算出する車両走行速度情報の信頼性を担保すること。	B
	AEBS-FSR-FEB-002	走行データ出力 ECU の自己診断を行うこと。	B
	AEBS-FSR-FEB-003	走行データ出力 ECU の自己診断結果を送信すること。	B
ステアリング 出力 ECU	AEBS-FSR-STE-001	ステアリング出力 ECU で算出する操舵角情報の信頼性を担保すること。	B
	AEBS-FSR-STE-002	ステアリング出力 ECU の自己診断を行うこと。	B
	AEBS-FSR-STE-003	ステアリング出力 ECU の自己診断結果を送信すること。	B
Planner (AEBS)	AEBS-FSR-PLN-001	障害物検出を行ってから、100ms 以内に緊急停車の実行を判断すること。	B
	AEBS-FSR-PLN-002	緊急停車判断を行ってから、100ms 以内に命令を送信すること。	B
	AEBS-FSR-PLN-003	Planner で算出する制御情報の信頼性を担保すること。	B
	AEBS-FSR-PLN-004	Fusion の自己診断結果を受信すること。	B
	AEBS-FSR-PLN-005	走行データ出力 ECU の自己診断結果を受信すること。	B
	AEBS-FSR-PLN-006	ステアリング出力 ECU の自己診断結果を受信すること。	B
	AEBS-FSR-PLN-007	Planner の自己診断を行うこと。	B
	AEBS-FSR-PLN-008	Planner の自己診断結果を送信すること。	B
	AEBS-FSR-PLN-009	Fusion が故障している場合、アクセル/ブレーキ制御 ECU への出力を停止すること。	B
	AEBS-FSR-PLN-010	Fusion からの出力がない場合、アクセル/ブレーキ制御 ECU への出力を停止すること。	B
	AEBS-FSR-PLN-011	走行データ出力 ECU が故障している場合、アクセル/ブレーキ制御 ECU への出力を停止すること。	B
	AEBS-FSR-PLN-012	走行データ出力 ECU からの出力がない場合、	B

		アクセル/ブレーキ制御 ECU への出力を停止すること。	
	AEBS-FSR-PLN-013	ステアリング出力 ECU が故障している場合、アクセル/ブレーキ制御 ECU への出力を停止すること。	B
	AEBS-FSR-PLN-014	ステアリング出力 ECU からの出力がない場合、アクセル/ブレーキ制御 ECU への出力を停止すること。	B
アクセル/ ブレーキ制御 ECU	AEBS-FSR-ACL-001	緊急停車を開始してから 3000ms 以内に停車すること。	B
	AEBS-FSR-ACL-002	アクセル/ブレーキ制御の信頼性を担保すること。	B
	AEBS-FSR-ACL-003	Planner の自己診断結果を受信すること。	B
	AEBS-FSR-ACL-004	アクセル/ブレーキ制御 ECU の自己診断を行うこと。	B
	AEBS-FSR-ACL-005	Planner が故障している場合、非常停車すること。	B
	AEBS-FSR-ACL-006	Planner からの出力がない場合、非常停車すること。	B
	AEBS-FSR-ACL-007	アクセル/ブレーキ制御 ECU が故障している場合、非常停車すること。	B

システム構成の AEBS に対して上記の機能安全要求を割り当てることで安全な設計となっているか、検証を実施した。各機能ブロックの誤動作について網羅的に安全分析を実施し、安全目標を満たす設計であることを確認した。

4.2.1.4. 認証機関による技術レビューならびに評価

作成した文書について、認証機関 SGS との技術ミーティングを開催し、レビューを受けた。

実施日時：2020 年 1 月 29 日 10 時 00 分～17 時 00 分

実施場所：株式会社ヴィッツ 204A 会議室

実施概要：認証機関 SGS による、以下のドキュメントのレビュー

- MaaS システム安全コンセプト.docx
- MaaS システム_ユースケース一覧.xlsx
- MaaS システム_分析表.xlsx
- AEBS 安全コンセプト.docx
- AEBS 安全分析.xlsx

レビューの総評を以下に示す。

SGS 総評

■ 出来ていること

- 前提条件が明確になっており、網羅性の高いコントロールストラクチャをベースに、一貫性のある安全コンセプトとなっている。

■ 改善点

- 重要事項の漏れはないが、効果確認の手法や設計検証における優先度/基準を検討頂くとよい。

3

すなわち、SGS との技術ミーティングでは、以下の結果となった。

- AI 搭載システムについて、サービスレベルから定義・分析を実施することにより、スコープが明確で説明性の高い安全コンセプトであることを認めてもらうことができた。
- コントロールストラクチャの網羅性を高め、抽出した UCA とユースケースシナリオを組み合わせることにより、実践的な STAMP/STPA となっていることを評価してもらうことができた。

4.2.2. 人工知能の種別・安全度・安全対策方法の分類（サブテーマ 2-3）

4.2.2.1. システム定義やハザード分析&リスクアセスメントの実施方法や観点の整理

MaaS 安全コンセプトの作成時に得た知見を基に、AI 搭載システムを用いたサービスにおける、システム定義やハザード分析&リスクアセスメントの実施方法、観点を整理した。

また、各種規格を調査し、以下のような観点を取り込んだ。

- ISO/PAS 21448：自動運転システム開発における留意点
- IEC 62853：多数のステークホルダが関係するサービスにおける、開発者の責任範囲
- IEC 60721：考慮すべき環境条件

システム定義ではサービスレベルのシステムのスコープを明らかにし、機能や制約、およびシステムの構成(初期のアーキテクチャ)を定義するため、以下の内容を記載することとした。

- システム概要
- 利用条件
システムを運用時に守るべき条件、条件が守られる根拠、条件が侵害された場合の対応
- 環境条件
システムの運用時の想定環境、環境が維持される根拠、環境が維持されない場合の対応
- ユースケース
システムのふるまいを確認する他、後述する非安全シナリオ導出で使用する異常系についても記載する
- 各サブシステムの機能

ハザード分析及びリスクアセスメントは、定義したシステムが安全である（＝許容できないリスクがない）ことを示すために実施する。

ハザード分析の手法として、コンポーネント（サブシステム）同士が相互に作用し、また外部環境を考慮するような複雑なシステムの分析に適しているとされる STAMP/STPA を選択した。STAMP/STPA の進め方やリスクアセスメントへの繋ぎ方は厳密に定められていない。本研究事業では既存の手順を参考にしつつ、シナリオベースでリスクアセスメントを実施するため、以下の手順を提案した。

➤ アクシデント・ハザード・安全制約の決定

まず、損失につながるようなイベントとなるアクシデントを定義する。損失として考えられるものは人命損失、けが、物損、環境汚染、経済的損失など様々だが、何のために何を分析するのかを明確にするため、アクシデントの範囲を定める必要がある。

次に、最悪な条件と重なることでアクシデントにつながるハザードを定義する。1つのアクシデントに対してハザードが1つになるとは限らない。また、2つ以上のハザードが同時に発生することにより初めてアクシデントとなる場合もある。

そして、ハザードを引き起こさない、すなわちシステムを安全に保つための要件となる安全制約を導く。

最後に、上記のアクシデント・ハザード・安全制約に対して ID を割り振り、アクシデント・ハザード・安全制約表としてまとめる。

➤ コントロールストラクチャの作成

安全を維持するための制御構造であり、コンポーネントとコントロールアクション、フィードバックの関係を現す、コントロールストラクチャを作成する。

コントロールストラクチャに示す範囲がそのままハザード分析の範囲となる。そのため、環境要件を考慮するシステムであれば、その環境要件もコントロールストラクチャに記載する必要がある。運搬、点検、修理など、運用中以外を分析範囲とする場合は、階層を分けて記載するとよい。

コントロールアクションは電子システムにおける入出力データとは必ずしも一致せず、コントロール対象のコンポーネントに対して振る舞いを行わせるものである。

➤ UCA の抽出

システムの振る舞いであるコントロールアクションが非安全に行われた時に起きる影響を、非安全動作（Unsafe Control Action、以降 UCA と記載）として抽出する。UCA の抽出にあたっては、以下の4つのガイドワードを用いる

- 与えられないとハザード
- 与えられるとハザード
- 早過ぎ、遅すぎ、誤順序でハザード
- 早過ぎる停止、長過ぎる適用でハザード

各コントロールアクションに対してガイドワードを適用し、それが安全制約を侵害するものであれば UCA として抽出し、UCA 一覧としてまとめる。

➤ UCA を含めたユースケースでの非安全シナリオ導出

ユースケースの基本フローに対して、コントロールアクションを行っている箇所に UCA を当てはめ、フローへの影響を非安全シナリオとして導出する。同じコントロールアクションに対して複数の UCA が当てはまる場合、それらは個別のシナリオとして記載する。

ここで明らかに不自然なシナリオとなり、定義したシステムでは発生し得ないと判断した場合は下記のパターンから対処方法を選択することができるが、その方針は統一しておく必要がある。

- 判断理由を確認できるようにした上で UCA 一覧から除外する
- 後述するリスクアセスメント時の発生頻度、曝露の確率などのパラメータを低下させる

➤ リスクアセスメント

各非安全シナリオについて、発生しやすいさ、発生した場合の被害がどの程度なのか等を定量的指標で示し、リスクの大きさを評価する。

想定するシステムによって適用するリスクアセスメントの手法は異なる。例として自動車向けの機能安全規格である ISO 26262-3:2018 では、曝露の確率(E)、コントローラビリティ(C)、シビアリティ(S)の3つのパラメータを用い、リスクの大きさをASILで評価する。

曝露の確率(E)

曝露は「分析対象の故障モードと組み合わせると、運用状況が危険になる可能性がある状態」と定義される。ISO/PAS 21448(SOTIF)においては故障モードではなく、性能限界や合理的に予見可能なミスユースを考慮する。さらに想定される運用状況の特性に応じ、その「期間」と「発生頻度」のどちらかの指標が用いられる。

曝露の確率に影響する例として、機械音声によるアナウンスがある。シナリオを分析することで、アクシデントに繋がる動作の前にアナウンスがあり、それを人間が聞くことによって事前に回避できることを確認すれば、曝露の確率は下がることになる。

コントローラビリティ(C)

コントローラビリティは「当事者のタイムリーな反応により、場合によっては外部方策の支援により、特定された危害または損害を回避する能力」と定義されている。つまり、アクシデントに繋がる状況が発生した際、人間またはシステムによる回避行動により、危害を回避または軽減できるかという確率がその指標になる。

曝露の確率とコントローラビリティはトレードオフの関係となる場合がある。例えば、アナウンスに従った行動を人間が取っていればアクシデントに繋がらないとする。この場合、アナウンスを無視、あるいは聞こえなかった場合は回避が難しいということになる。

システムによる回避行動の例として、自動運転システムの緊急停車機能がある。障害物が車両の前方に出現、あるいは白線逸脱を検出できなかったとしても、障害物検出による緊急停車ブレーキが作動すれば、アクシデントを回避できると考えられる。ただし、回避行動にあたる機能は、曝露した事象と独立している必要がある。上記の例では、障害物検出機能やブレーキ制御が故障する事象だと緊急停車機能によりアクシデントを回避することができない。

シビアリティ(S)

シビアリティは「潜在的な危険事象における、一人以上の個人に対して発生する可能性のある危害程度の見積もり」と定義されている。アクシデントが発生した際の、人に対する危害の大きさの程度のことである。

この時、人間またはシステムによる回避行動を考慮することができる。例えば自動運転車両が障害物に衝突するアクシデントの場合、障害物検出による緊急停車ブレーキが作動することにより、衝突時の速度は通常走行時に比べて下がっていることが考えられる。衝突時の速度が20km/h未満であれば、車両内の乗客に怪我はない、または軽症で済むと判断できる。

4.2.2.2. 人工知能搭載システムの安全論証手法の分類と具体的な手順の整理

AIは、出力に一定の誤りが生じる、入出力の整合確認が困難な場合がある、内部設計の二重化ができない、などといった特性を持つ。

上記のAIの特性を考慮した上でAI搭載システムに対して機能安全設計を行う方法について検討した。その結果2つの方針、計7パターンを提案した。

➤ 方針1：AI自体を安全設計する

➤ 1a) 機能安全開発パターン

AIを機能安全規格に準拠したプロセスで開発することにより、AIが十分に高い信頼性があることを示す方法である。このような機能安全プロセス相当の実現が期待できる技術として、「説明可能なAI(eXplainable AI; XAI)」があるが、現時点では有用な手法は見出されてい

（令和元年度戦略的基盤技術高度化支援事業）

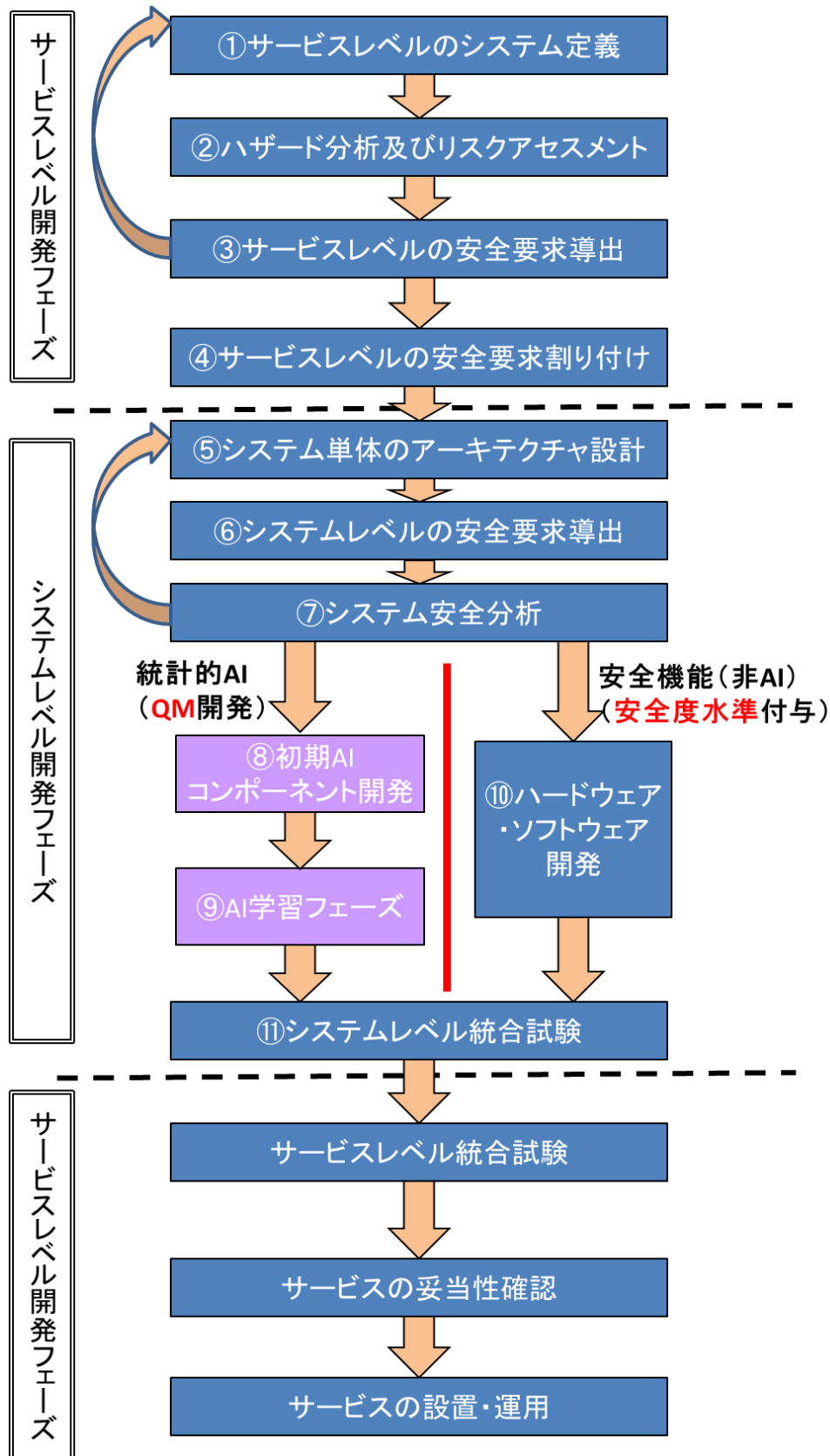
「自律的自動運転の実現を支える人工知能搭載システムの安全性立証技術の研究開発」研究成果報告

ない。

- 1b) 安全性評価パターン
AI について機能安全規格適合相当の安全性評価を行う手法である。ブラックボックス評価手法の位置づけになる。有力な手法として、センサシステムの機能安全規格 IEC 62998 の評価方法が適用可能だと考える。
- 1c) Proven-in-use パターン
AI について、機能安全規格が要求する Proven-in-use（使用実績）基準を示すことによって、十分な安全性があると評価する方法である。ブラックボックス評価手法の位置づけになる。例えば、IEC 61508-7:2010 C.2.10.1 に記載されている基準を満たすことで、対象コンポーネントの Proven-in-use を主張することができる。
- 1d) 多重化設計パターン
QM 品質相当のコンポーネントを多重化（冗長）設計することにより、安全度水準を満たす方法である。例えば、ISO 26262 では最大二重故障までしか考慮する必要がないため、三重故障は safe fault として扱ってよい。
三重化設計といっても、多数決ではなく、冗長化である点に注意されたい。
- 方針 2：AI は安全設計せず、外部に安全メカニズムを設ける
 - 2a) 機能安全機能パターン
QM 品質の AI とは独立した機能安全機能（SIL）を搭載することによって、AI が誤った結果を出力したとしても危険にならない設計にする方法である。この設計パターンは、安全機能による監視の強度によっては、逆に AI のよさを損なう可能性もある。
 - 2b) 比較パターン
QM 品質の AI と機能安全品質の他のコンポーネント（SIL）による多重化を行い、出力結果を比較する方法である。AI の出力結果が安全であれば、その結果を用い、もし、安全でなければ、他のコンポーネント（SIL）の結果を用いる。
 - 2c) 防御設計パターン
動的にロバスト性の確認をするための防御設計（ガード）を搭載する方法である。ロバスト性とは、あるシステムが応力や環境の変化といった外乱の影響によって変化することを阻止する内部的な仕組みのことである。ガードに引っ掛かった場合、停止制御によってシステムの安全を担保する。
2a) 機能安全機能パターンとの違いは、AI に近い箇所での簡易的なガード処理を行う点である。2b) 比較パターンとの違いは、異常を検知して停止制御をする点である。

4.2.2.3. 人工知能搭載システムの安全説明力の高い開発プロセスの整理

4.2.2.2の提案のうち、方針 2 に基づいた安全開発プロセスを考案、整理した。下図にプロセスの全体像を示す。



① サービスレベルのシステム定義
4.2.2.1を参照とする。

②ハザード分析及びリスクアセスメント

4.2.2.1を参照とする。

③サービスレベルの安全要求導出

リスクアセスメントの結果から、許容できない非安全シナリオを持つUCAに対する安全方策を検討し、安全要求として導出する。

④サービスレベルの安全要求割り付け

導出した安全要求を各サブシステムに割り付け、それらの安全目標とする。

⑤システム単体のアーキテクチャ設計

次項で導出するシステムレベルの安全要求を配置するため、各サブシステム(システム単体)の機能ブロック図を作成する。なお、AIコンポーネントには安全要求を割り当てない方針とするため、方針2の2a)～2c)のいずれかのパターンに合致するような機能ブロック図とする。

⑥システムレベルの安全要求導出

システムに割り付けられた安全目標を分解し、安全要求を導出する。この時、機能ブロックの故障に関する安全要求についても検討する。安全目標は各機能ブロックに割り振ることのできる単位まで、階層的に分解する。

⑦システム安全分析

各機能ブロックの誤動作について網羅的に安全分析を実施し、安全目標を満たす設計であることを確認する。

⑧初期AIコンポーネント開発

AI学習前のアルゴリズム部分について開発を行う。AIコンポーネントには機能安全要求を割り当てない方針としているため、ISO 9001等の品質マネジメントシステムに従って開発を行う。

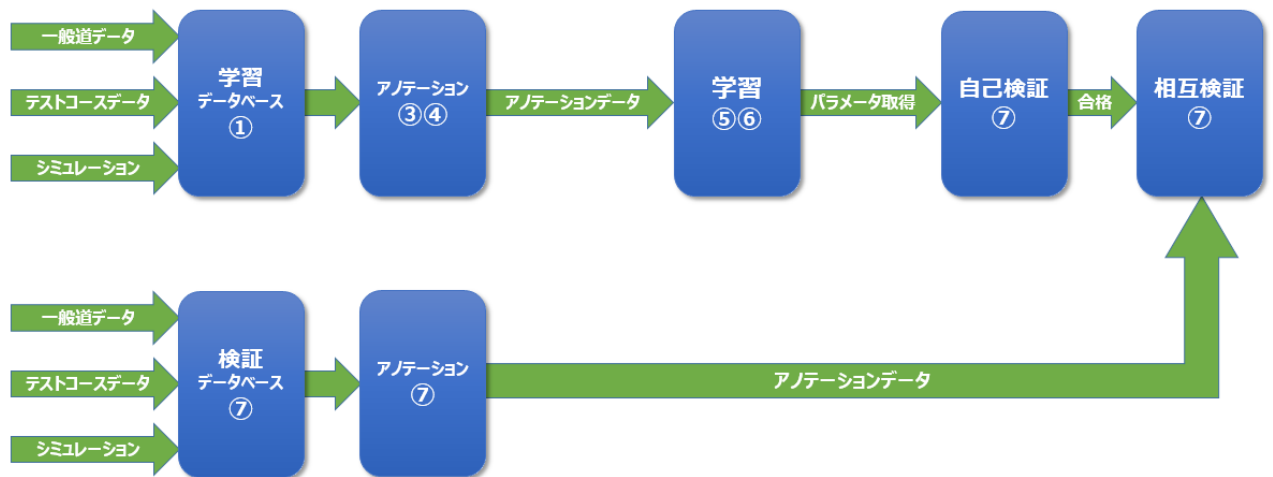
⑨AI学習フェーズ

AI学習フェーズにおいても、初期AIコンポーネント開発と同様に、ISO 9001等の品質マネジメントシステムに従って開発を行う。

本研究事業では自動運転システム向けのAI学習フェーズとして、昨年度の研究において、Automotive SPICEのテーラリングを行ったプロセスを考案した。今年度はこれに加え、SOTIFに記載されたAI学習フェーズのワークフロー、説明可能性の高いAIモデルワークフローの手法についても調査を行い、AI学習フェーズのプロセスとして整理した。

➤ AI学習フェーズのワークフロー(SOTIF)

SOTIF(ISO/PAS 21448) Annex Gの記載を基にしたAI学習フェーズのワークフローを以下に示す。



上記のうち「学習データベース」「アノテーション」は、昨年度考案した、Automotive SPICE のテーラリングを行ったプロセスの「学習データの収集」と、「学習」は「学習の実行」と、「自己検証」「相互検証」は「検証」と、それぞれ対応している。

また、上記図中の番号①～⑦は、次項「説明可能性の高い AI モデルワークフロー」と対応している。なお、②ベンチマークや再現性の確立は、ワークフロー全体に関連するプロセスである。

－ 学習データベース

一般道、テストコース、シミュレーション環境などからデータを収集し、学習用のデータベースを構築する。

－ アノテーション

収集したデータにタグ付けし、アノテーションデータを作成する。

－ 学習

アノテーションデータを AI に読み込ませることで学習を行い、AI の内部パラメータを決定する。

－ 自己検証

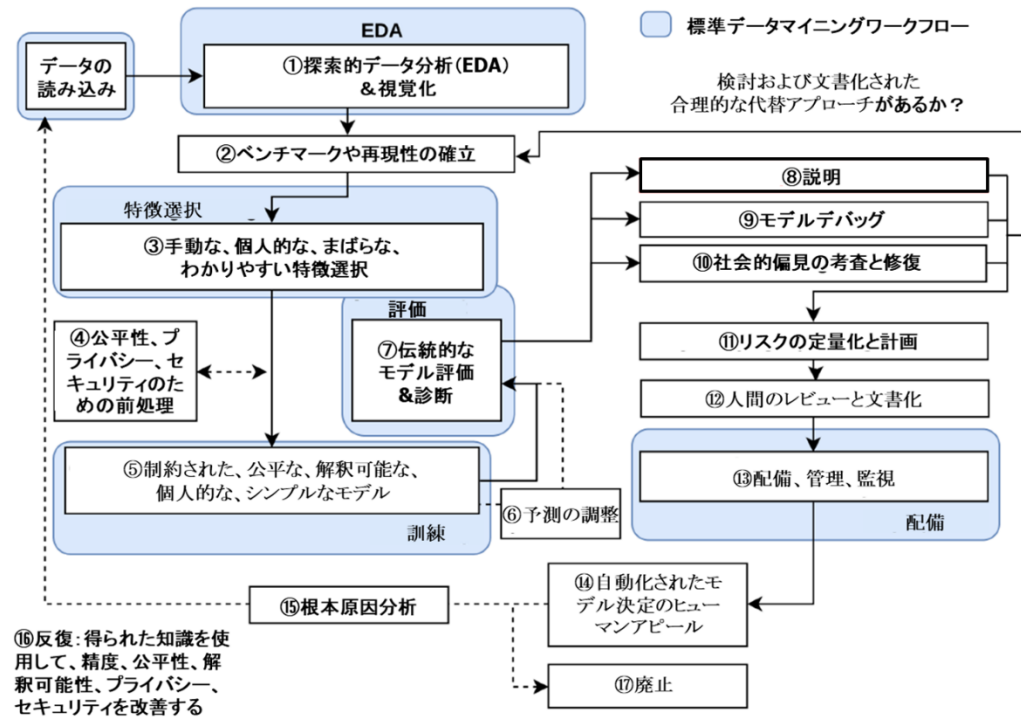
学習を行った AI の検証のため、アノテーション前のデータを AI に読み込ませ、その結果がアノテーションデータと一致していることを確認する。

－ 相互検証

学習に使用したものとは異なる検証データベースを使用し、相互検証を行う。

➤ 説明可能性の高い AI モデルワークフロー

説明可能性の高い機械学習（AI）モデルの開発ワークフローを以下に示す。



参考文献 : Proposed Guidelines for the Responsible Use of Explainable Machine Learning
<https://arxiv.org/pdf/1906.03533.pdf>

標準データマイニングワークフローに加え、さまざまな活動を実施することで、AI モデルの決定根拠（エビデンス）を押さえることができ、「AI モデルの開発プロセスに対する説明可能性」を高めることができる。これにより、機械学習における不透明さ、不正確さ、社会的バイアス、またはセキュリティの脆弱性によってもたらされる共通リスクに対する技術的解決策になる可能性がある。

ただし、本ワークフローは「AI モデルそのものをホワイトボックス化」する技術ではない。そのため、機能安全で要求される品質保証レベルを満たすには必要ではあるが、十分ではないと考える。

⑩ハードウェア・ソフトウェア開発

ハードウェアおよび安全関連系ソフトウェアそれぞれに対して、機能安全開発プロセスに従った開発を行う。システムレベルの設計から展開された要件をそれぞれ、ハードウェア要件、ソフトウェア要件に分解し、それぞれの設計を行う。ハードウェア開発、ソフトウェア開発では、それぞれの開発において統合テストまで実施する。

⑪システムレベル統合試験

すべてのハードウェア、ソフトウェアを統合し、システム全体での統合テストを実施する。この統合には AI コンポーネントも含まれる。実装された機能と性能に対して、検証と妥当性の確認を行う。さらに、AI に関連して統合テストとして、以下の確認を行う。

- AI と安全関連系のパーティショニングの検証
- 故障注入テスト（フォールトインジェクション）の実施
- 意図的に AI の故障や誤りを発生させてテストを実施する

4.2.2.4. 人工知能搭載システム安全ガイドラインの对外紹介の実施

AI プロダクト品質保証コンソーシアム（QA4AI）や人工知能に関する標準化委員会である ISO/IEC JTC1 SC42 と連携し、以下のセミナーによる技術紹介を実施した。受講者からは高い関心・評価を得ることができ、当研究企業の事業化に貢献する活動となった。

開催日	セミナー名	タイトル
2019 年 5 月 17 日	Open QA4AI Conference	自律的自動運転の実現を支える人工知能搭載システムの安全性立証技術の研究開発
2019 年 10 月 9 日	情報処理学会 短期集中セミナー	AI の安全性に関する標準化動向と安全開発手法の紹介

4.2.3. 令和元年度の活動のまとめ

以上の令和元年度の活動の結果、我々の安全設計の取り組みの方向性は正しいことを確認できた。また、以下の取り組みが本研究の強みであることを認識できた。

＜本研究の強み（サブテーマ 2 範囲）＞

- ・ 具体的な事例（自動運転レベル 4 システム）に適用した成果物を有する
- ・ AI 搭載システムの安全性論証パターンを整理し提案した唯一の研究
- ・ 安全説明力の高い学習プロセスを整理。特に AutomotiveSPICE のテラリング適用は唯一の研究

今後は、当研究成果を実開発に適用しながら、有効性を評価し、さらなる改善に努めていく。また、当研究成果の一部を業界で共有しながら、標準化に向けた普及活動に努めていく。

4.3. 人工知能搭載システムの開発コスト対応（サブテーマ 3）

4.3.1. 自動走行車両による実証実験

4.3.1.1. 実証実験 ACC システム

4.3.1.2. 設計

実証実験システムの構成図を図 3-3-1-1-1 に示す。基本的に昨年度に決定した構成と同様であるが、膨大なログデータを記録するためにノート PC および DrivePX2 にそれぞれ SSD を追加した。

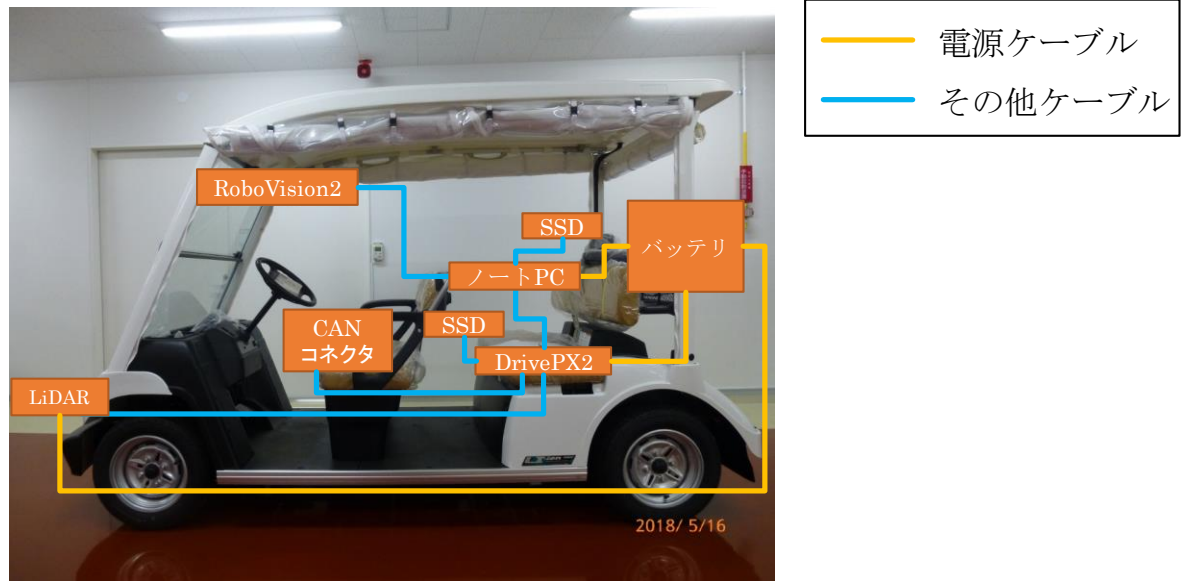


図 実証実験システム構成図

実証実験 ACC システムのソフトウェア構成図を示す。

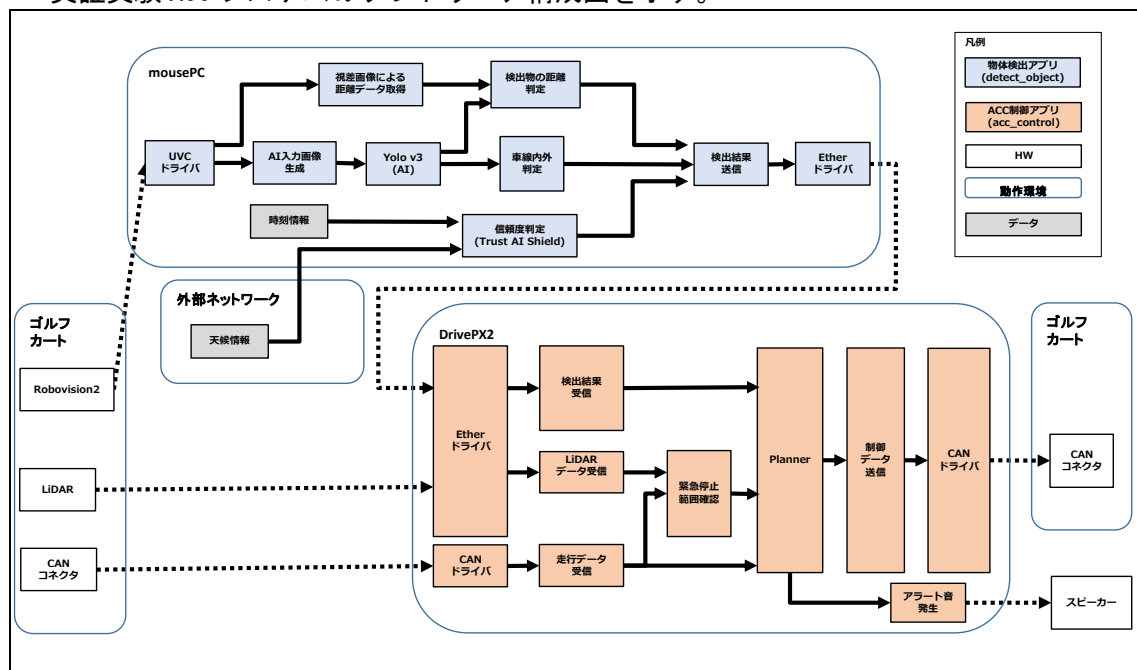


図 実証実験システムソフトウェア構成図

それぞれのソフトウェアコンポーネントの役割を表にまとめる。

表 ソフトウェアコンポーネントの機能

分類	ソフトウェアコンポーネント	機能
物体検出アプリ	AI 入力画像生成	Robovision2 から取得できる画像はステレオ画像である。検出物の距離判定を実施しやすくするため、距離データと同一範囲になるようにトリミングと収縮を実施。
	視差画像による距離データ取得	Robovision2 の API を使用して、視差画像から取得した距離データを取得。
	YOLO v3 (AI)	物体検出を行う AI アプリケーション。
	検出物の距離判定	物体検出 AI から取得した検出座標内に該当する距離データから最頻値を算出する。最頻値の値を検出物の距離として出力する。
	車線内外判定	事前に定めた車線幅にあたる座標をもとに、検出物が車線内か車線外かを判定する。
	信頼度判定 (Trust AI Shield)	天候および時刻情報を基に、物体検出 AI の出力の信頼度を判定する。
	検出結果送信	物体検出アプリの検出結果を ACC 制御アプリに UDP で送信する。
ACC 制御アプリ	検出結果受信	物体検出アプリから検出結果を受信する。
	LiDAR データ受信	LiDAR のデータを受信する。
	走行データ受信	ゴルフカートから走行データを受信する。
	緊急停止範囲確認	LiDAR データを参照し、緊急停止が必要な範囲内に物体がいるか否かを判定する。
	Planner	目標車速およびブレーキ指示を決定する。
	制御データ送信	Planner で決定した制御データをゴルフカートに送信する。
	アラート音発生	物体検出結果の信頼度が閾値を下回った場合に、アラート音を発生させる。

上記表に記載した機能に加えて、「Trust AI Shield」の機能の一部としてログの保存を実施する。それぞれのアプリケーションで保存するログデータを表に示す。

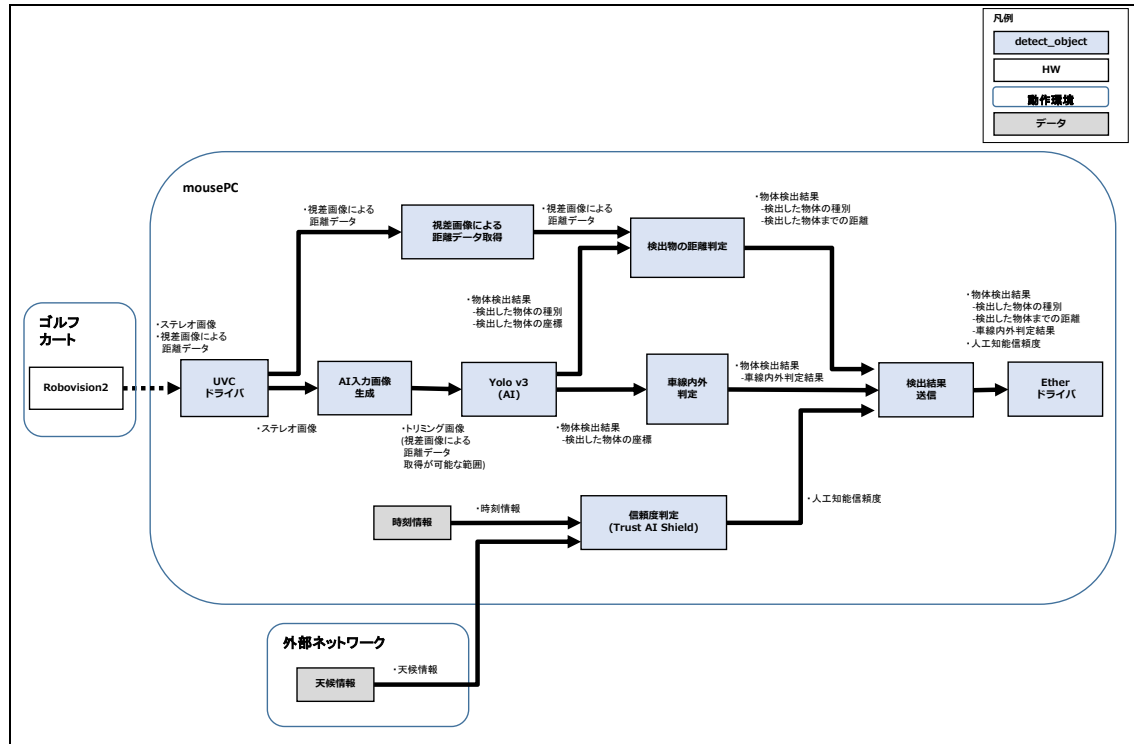
表 ソフトウェアコンポーネントの機能

分類	保存するデータ
物体検出アプリ	画像データ (ステレオ画像、トリミング&収縮後画像)
	視差画像から取得した距離データ
	物体検出結果 (検出した種別、検出した物体までの距離、車線内外判定結果)
	AI 信頼度
ACC 制御アプリ	LiDAR データ
	CAN 受信データ
	受信した物体検出結果 (検出した種別、検出した物体までの距離、車線内外判定結果)
	緊急停止フラグ
	制御データ (目標車速、ブレーキステータス)

(令和元年度戦略的基盤技術高度化支援事業)
「自律的自動運転の実現を支える人工知能搭載システムの安全性立証技術の研究開発」研究成果報告

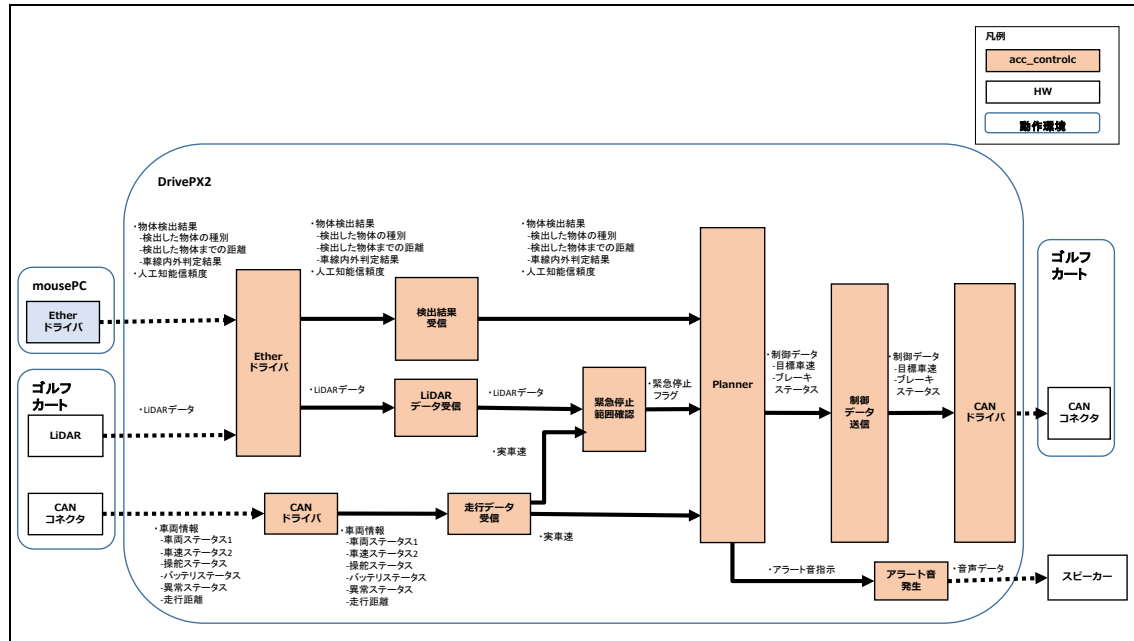
	CAN 送信データ
	アラート音指示

「物体検出アプリ」のデータフローを図にします。



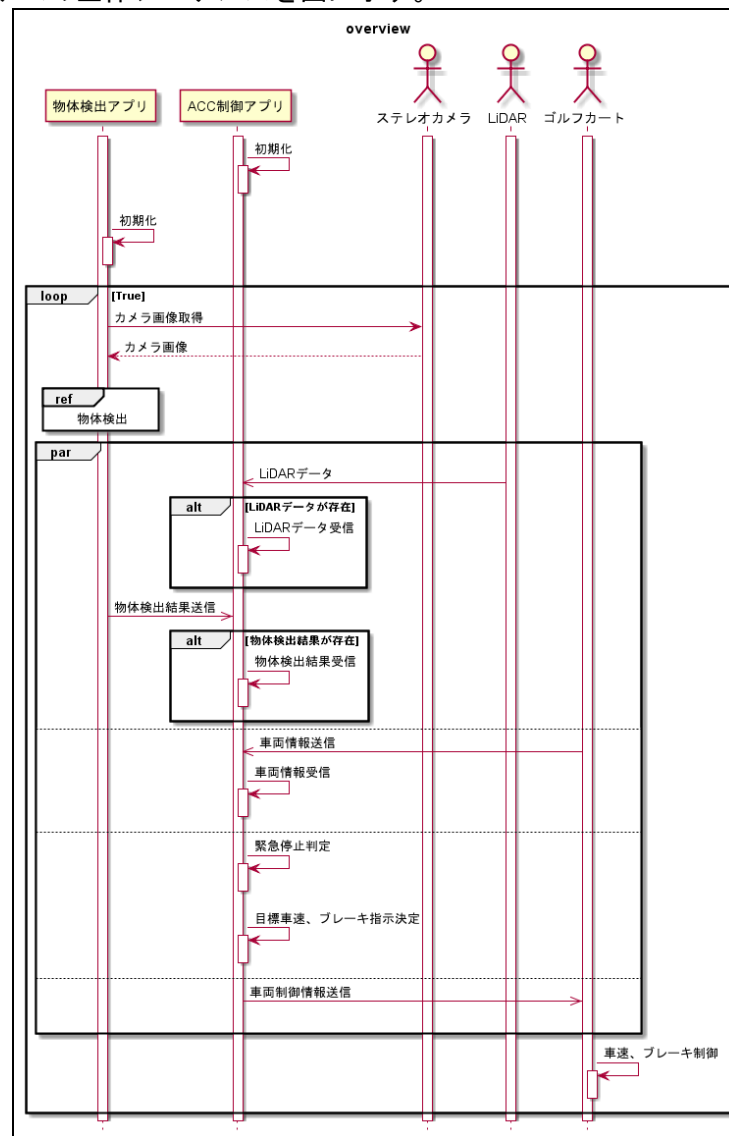
図「物体検出アプリ」データフロー

「ACC 制御アプリ」のデータフローを図に示す。



図「ACC 制御アプリ」データフロー

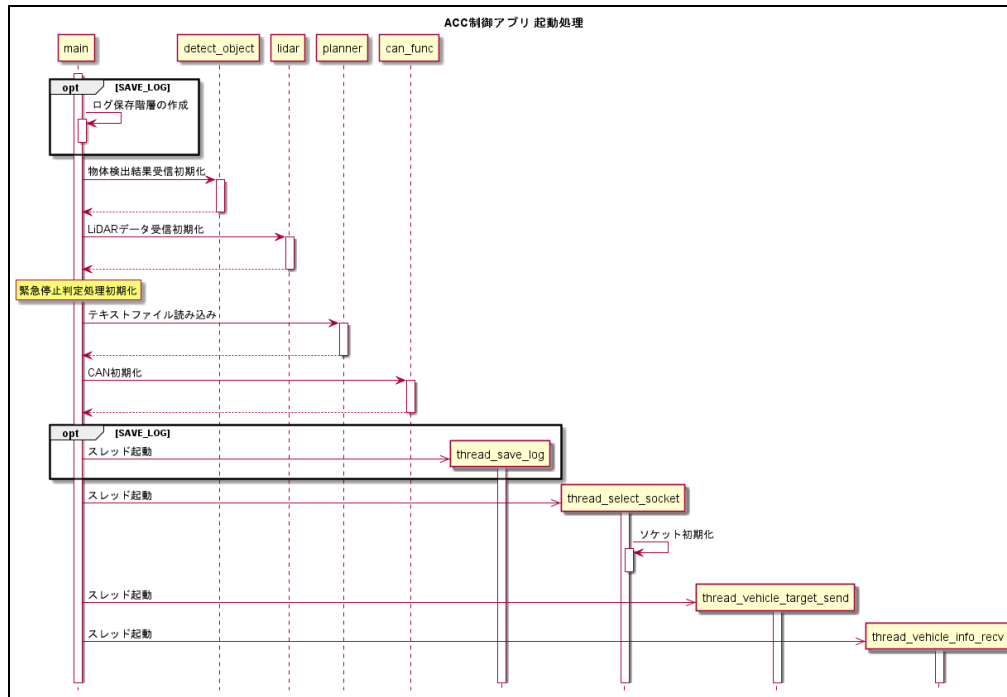
実証実験 ACC システムの全体シーケンスを図に示す。



図「実証実験 ACC システム」全体シーケンス

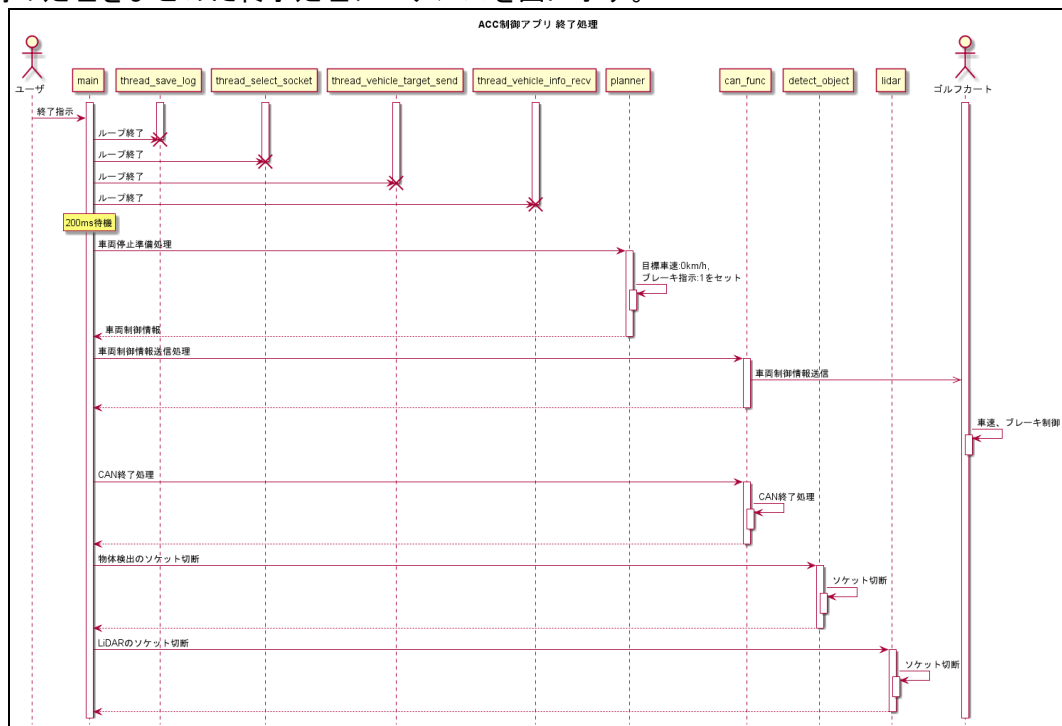
「ACC 制御アプリ」の詳細なシーケンス図を示す。「ACC 制御アプリ」の起動時の処理をまとめた起動処理シーケンスを図にしめす。

(令和元年度戦略的基盤技術高度化支援事業)
「自律的自動運転の実現を支える人工知能搭載システムの安全性立証技術の研究開発」研究成果報告



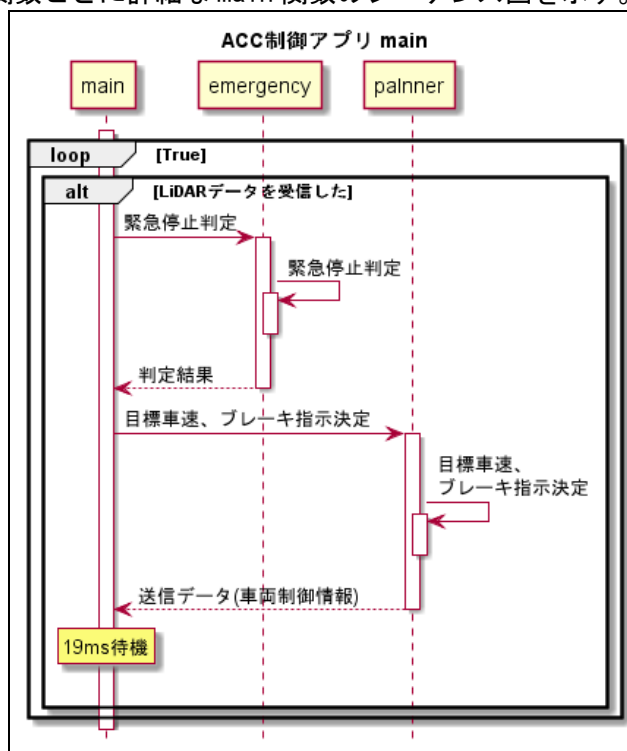
図「ACC 制御アプリ」起動処理シーケンス

終了時の処理をまとめた終了処理シーケンスを図に示す。



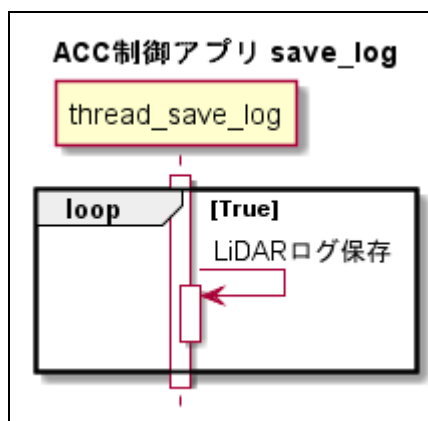
図「ACC 制御アプリ」終了処理シーケンス

「ACC 制御アプリ」の関数ごとに詳細な main 関数のシーケンス図を示す。

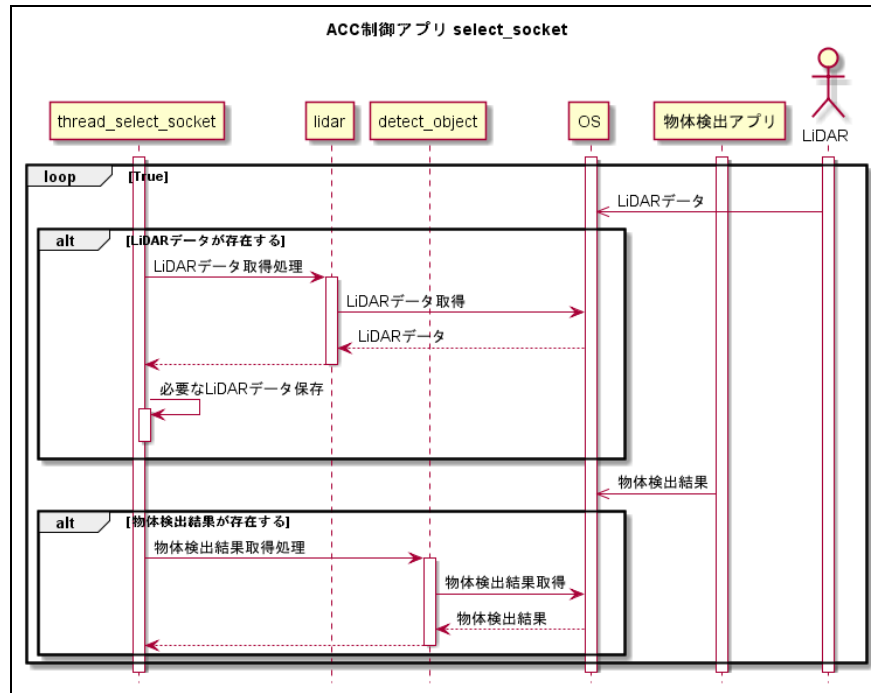


図「ACC 制御アプリ」main 関数シーケンス

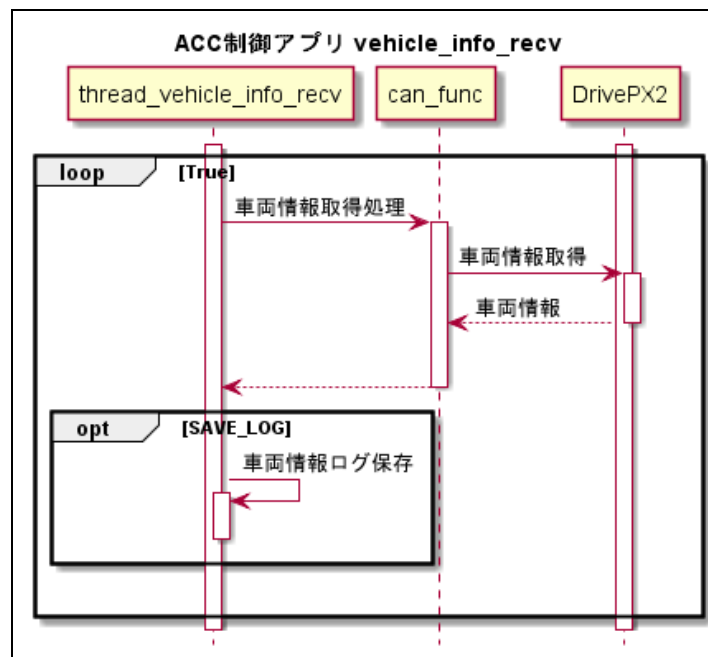
以下の図群は thread 関数であり、並列に処理を実施する関数のシーケンスである。



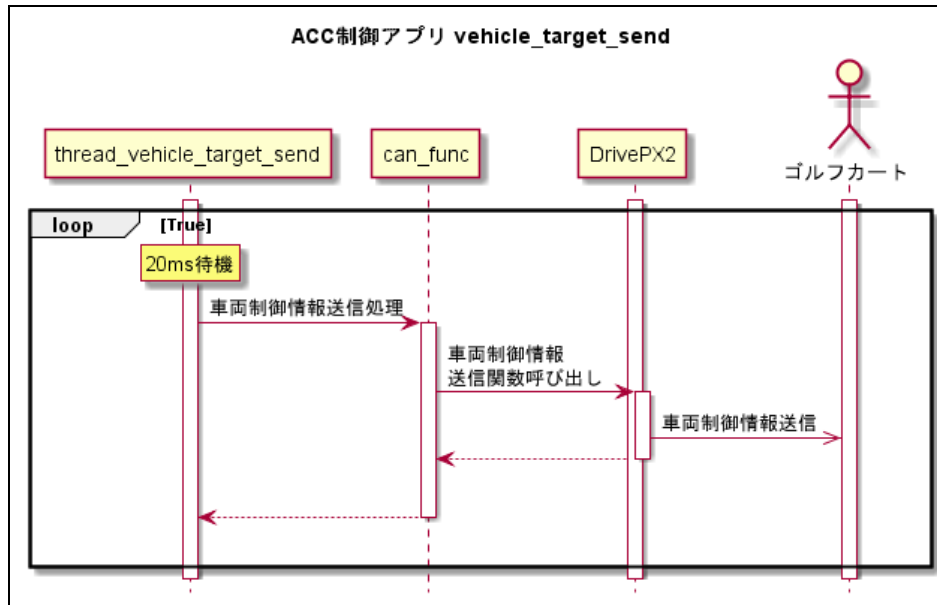
図「ACC 制御アプリ」save_log シーケンス



図「ACC 制御アプリ」select_socket シーケンス

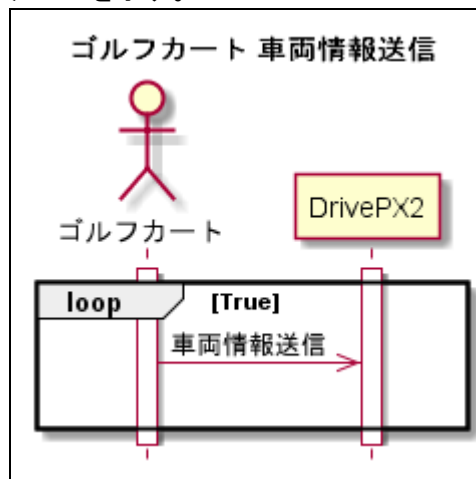


図「ACC 制御アプリ」vehicle_info_recv シーケンス



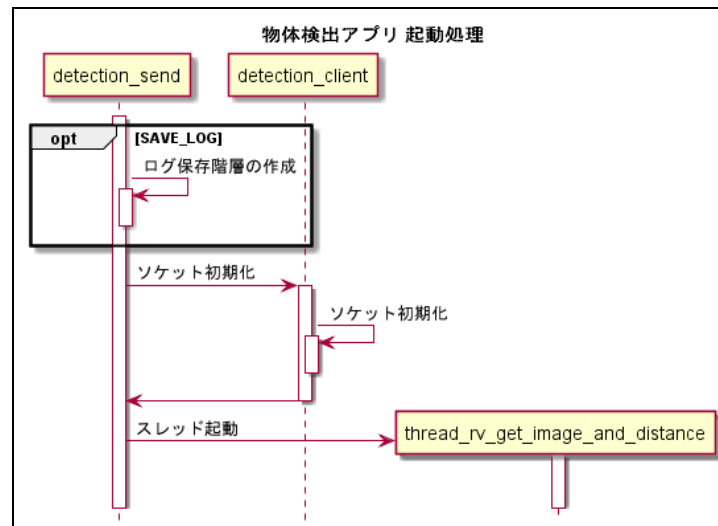
図「ACC 制御アプリ」 vehicle_target_send シーケンス

図は車両情報送信部のシーケンスを示す。



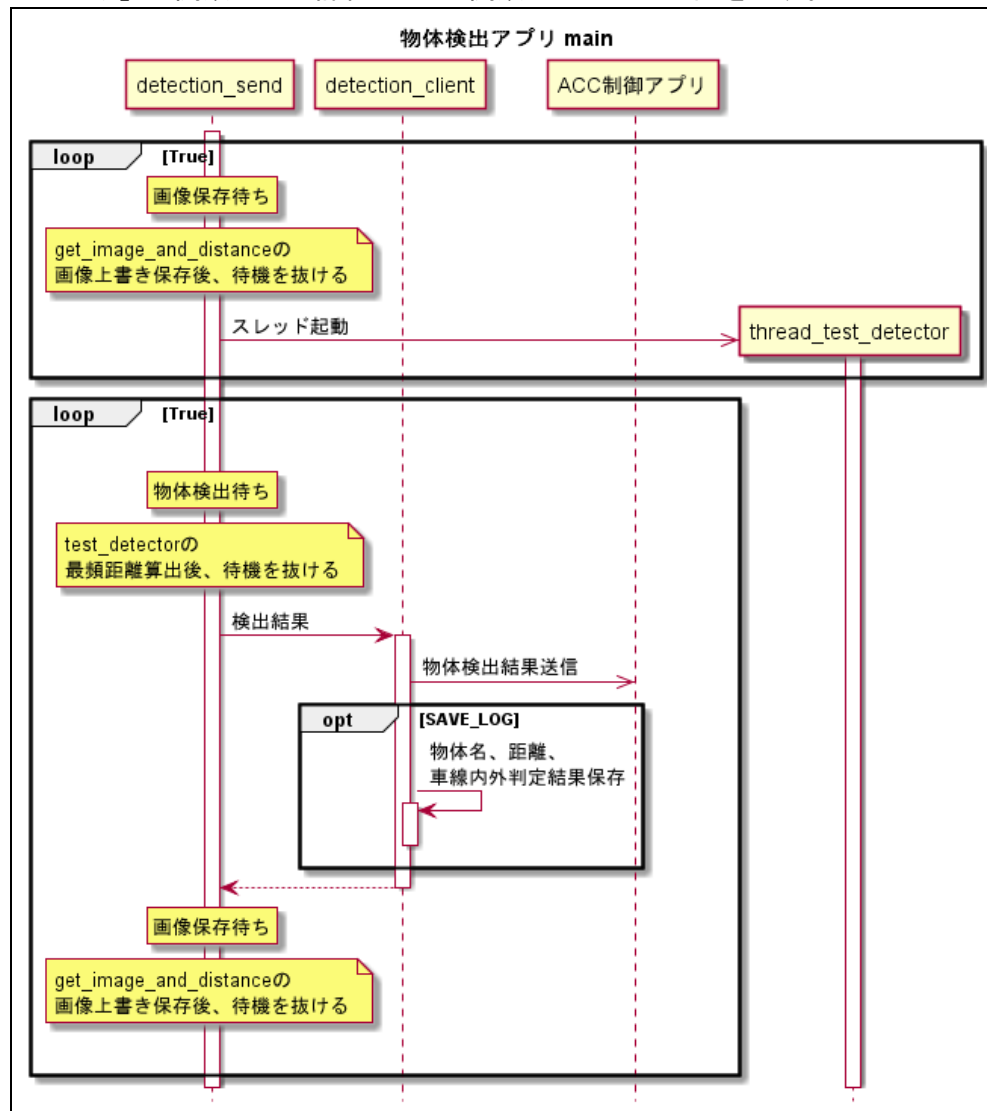
図「ACC 制御アプリ」車両情報送信シーケンス

次に、「物体検出アプリ」の詳細なシーケンス図を示す。「物体検出アプリ」の起動処理シーケンスを図に示す。



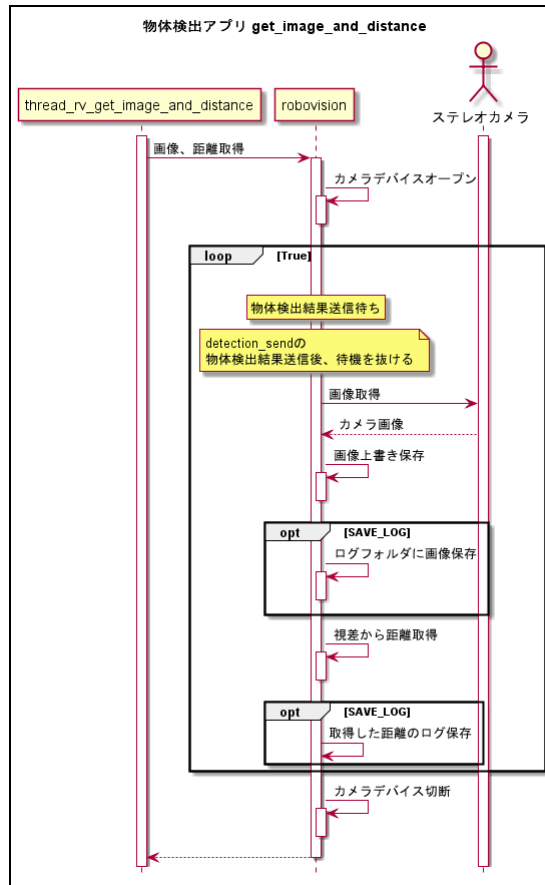
図「物体検出アプリ」起動処理シーケンス

「物体検出アプリ」の関数ごとに詳細な main 関数のシーケンス図を示す。

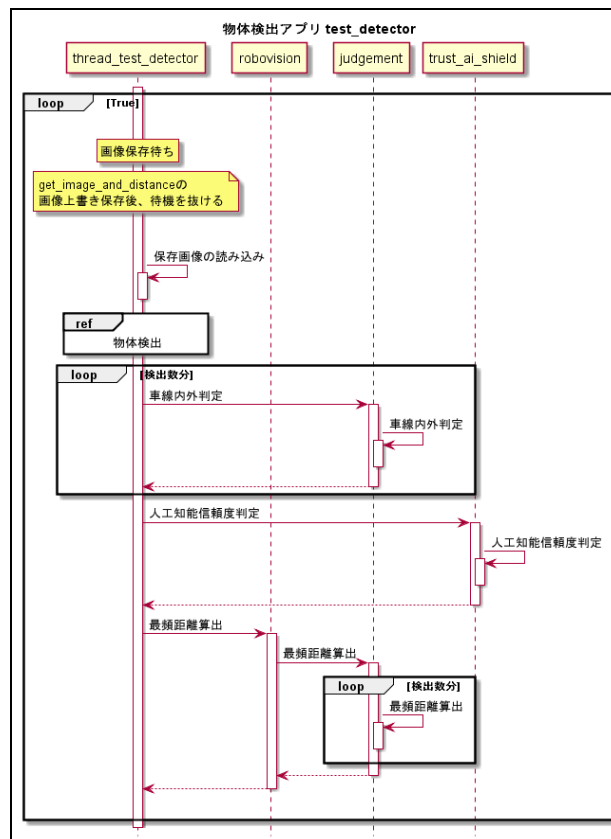


図「物体検出アプリ」main 関数シーケンス

以下図群は thread 関数であり、並列に処理を実施する関数のシーケンスである。



図「物体検出アプリ」get_image_and_distance 関数シーケンス



図「物体検出アプリ」test_detector 関数シーケンス

4.3.1.3. (3-3-1-2) 評価

実証実験は仮想環境および実機環境の2つで実施した。シミュレータ環境には図 3-3-1-2-1 に示す自動運転仮想検証システム ViViD を用いた。実機環境でのテストを実施する前段階として、作成したソフトウェアが意図通りに動作することを ViViD で確認した。



図 ViViD 画面

実機環境での実証実験は図に示す名古屋大学のシャーシダイナモ施設を利用して実施した。実機環境としてシャーシダイナモ施設を選択した理由は、車両がダイナモローラ上を走行するため、車両が想定外の動作を行った場合でも安全に評価を実施することができるためである。ただし、シャーシダイナモ施設は車両周辺の空間が実際の道路環境よりも狭く、壁などの障害物も多い。そのため、実証実験 ACC システムの一部のパラメータをシャーシダイナモ用に調整した上で実験を実施した。



図 シャーシダイナモ施設での実証実験風景

シャーシダイナモで実施した動作確認テストおよび走行テストを、表に示す。これらのテストを実施することで、実機環境で ACC システムが問題なく動作することを確認した。

表 実証実験確認テスト項目

大項目	中項目	実施件数
動作確認テスト	CAN 送信での実機動作確認	31
	安全機能の動作確認	3
走行テスト	障害物がない場合の動作確認	1
	車線内に障害物がある場合の動作確認	27
	車線外に障害物がある場合の動作確認	6
	緊急停止範囲内に障害物がある場合の動作確認	5

図は走行テストを行った際に、シャーシダイナモ施設で測定した車速および距離の概算を測定した結果である。横軸は時間であり、単位は秒である。

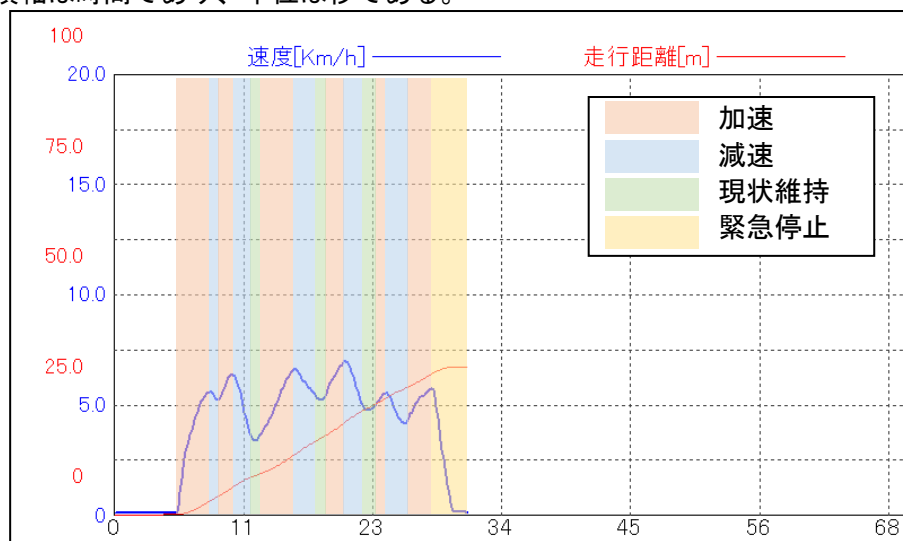


図 実証実験時の車速測定結果

自動走行車両が加速、減速および現状維持で速度の調整を行っていることが分かる。31 秒頃からは緊急停止で急激に減速し停車した。これは AI が誤検出・誤判断をしたことで、危険なエリアに物体が入ったことを LiDAR が検知したためである。実証実験システムの ACC 機能と安全機能の両方が動作していることを、以下の実験結果から確認できた。

- ・減速、加速、現状維持で速度の調整を行っていた
- ・危険なエリアに物体を検知したので緊急停止を実施し停車した

シャーシダイナモ施設は屋内施設であるが、本来の想定環境は屋外であるため、屋外でのセンサ認識結果の確認を実施した。物体検出 AI の出力画像を図に示す。



図 屋外でのセンサ認識結果

車両パネル、実際に駐車されている車両および人を過不足なく検出しており、物体検出 AI が問題なく物体を検出することを確認できた。

4.3.2. 「Trust AI Shield」

4.3.2.1. 「Trust AI Shield」検討結果

「Trust AI Shield」は AI の誤検出・誤判断を検出する、または AI の出力結果の信頼性を判断するソフトウェアコンポーネントである。精度が安定しない AI と「Trust AI Shield」を一緒に使うことで、AI の利益を活かしながら信頼性も高めることができる。

表にカメラを使用する AI に影響を与える可能性があるファクターリストを示す。このファクターリストは、AI 搭載システムの評価項目の作成に転用できることを想定している。

表 ファクターリスト

分類	ファクター	分類	ファクター
天候	快晴	フレームの連続性	想定し得るもの
	晴		想定外のもの
	くもり	検出物のカテゴリー	想定し得るもの
	雨		想定外のもの
	大雨	他のセンサとの整合性	一致
	雷		不一致
	雪		部分的に一致
	みぞれ	方角	東
	ひょう		西
	あられ		南
	霧		北
	台風	場所	舗装された道路
時間	早朝		未舗装の道路
	朝		トンネル
	昼	フレームレート	-
	夕方	仕様	画角
	夜		解像度
地域	-		ピント(焦点距離)
気温	-		
季節	春		
	夏		
	秋		
	冬		

本研究の実証実験システムでは、ファクターリストのうち天候および時刻を利用して「Trust AI Shield」の設計および実装を行った。

システム動作時のログを保存する機構を持つことも「Trust AI Shield」の機能としている。これはログを保存しておくことで、不具合が発生した際に原因を解析や対策をすることができるため、AI 搭載システムの保守に役立てることができるからである。

4.3.2.2. 「Trust AI Shield」のコンポーネントパターン検討案

「Trust AI Shield」のコンポーネントパターンを検討した結果を本章で示す。ただし、本章で検討した結果は実証実験システムには反映していない。今後の AI 搭載システムに対する「Trust AI Shield」を設計・実装する際の参考になる成果として記載する。

「Trust AI Shield」のコンポーネントのパターンとして、以下の表群の3つを想定する。それぞれのパターンを図群に示す。

表 コンポーネントパターンまとめ

パターン名称	概要
(1) AI 入出力監視パターン	AI の入力と出力の情報のみから AI の誤検出・誤判断を検出する/信頼度を判定する
(2) AI 出力監視パターン	AI の出力および AI 以外の情報を用いて AI の誤検出・誤判断を検出する/信頼度を判定する
(3) AI 入出力監視および AI 出力監視組合せパターン	(1) (2) のパターンの両方を実施し AI の誤検出・誤判断を検出する/信頼度を判定する

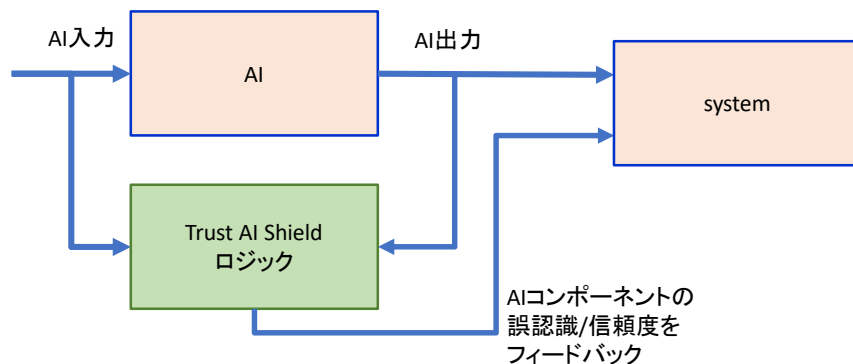


図 AI 入出力監視パターンイメージ

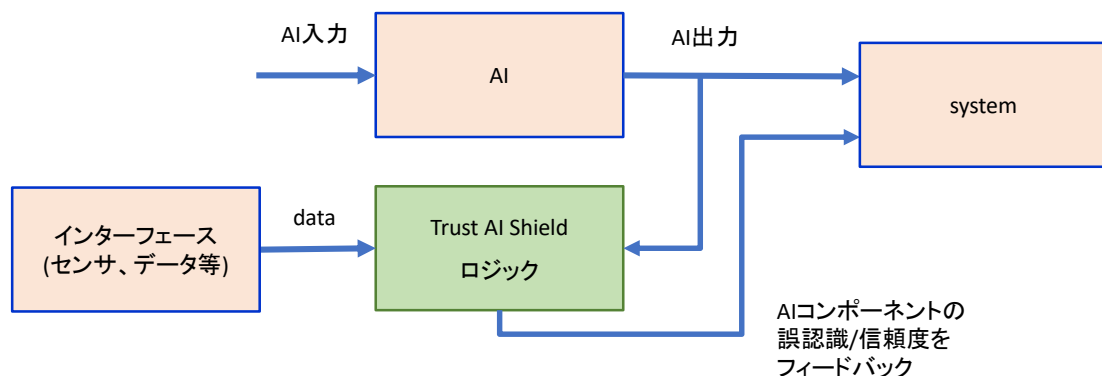


図 AI 出力監視パターンイメージ

（令和元年度戦略的基盤技術高度化支援事業）
「自律的自動運転の実現を支える人工知能搭載システムの安全性立証技術の研究開発」研究成果報告

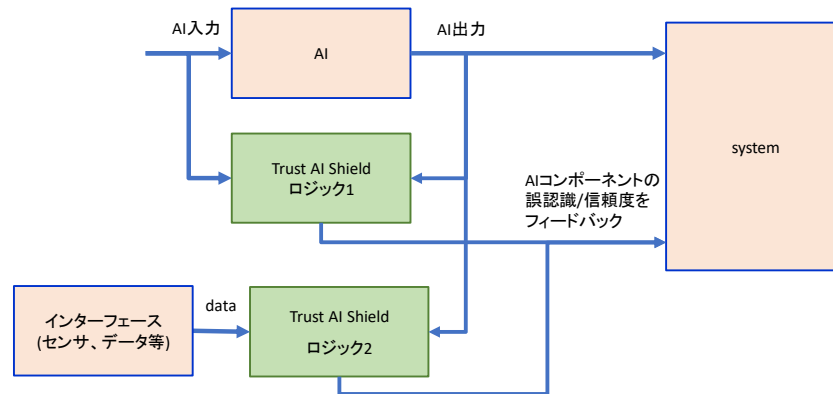


図 AI 入出力監視および AI 出力監視組合せパターンイメージ

上記に示したイメージは抽象的なものであり、より具体的なシステムを想定して「Trust AI Shield」のコンポーネント案を検討した。

物体検出 AI 単一設計での「Trust AI Shield」モジュール案を表にします。

表物体検出 AI 単一設計での「Trust AI Shield」モジュール案

No.	項目	説明	ユースケース
1	画像真ん中付近で検出した人が次のフレームで検出しなくなる	画像の真ん中付近で検出した人が、次のフレームで写りこまないのは人の移動速度を超えているため異常と判断可能	①画像真ん中で人を検出したフレームが誤検出 ②次のフレームで検出しなかったフレームが誤検出
2	前のフレームで人を検出していなかったにも関わらず、次のフレームで人を検出する	前のフレームで写りこんでいないのにも関わらず、画像の真ん中付近で人を検出するのは人の移動速度を超えているため異常と判断可能	①前のフレームで人を検出しなかったフレームが誤検出 ②次のフレームで人を検出したフレームが誤検出
3	入力に差分が見られないにも関わらず、次のフレームで車両や人を検出する	入力画像に差分がないのにも関わらず、新しい物体が存在することは異常と判断可能 ※ただし、車両が走行することにより発生する背景の差分と区別するためには何らかの工夫が必要	①前のフレームで人を検出しなかったフレームが誤検出 ②次のフレームで人を検出したフレームが誤検出
4	入力に差分が見られないにも関わらず、次のフレームで車両や人を検出しない	入力画像に差分がないのにも関わらず、物体が消失することは異常と判断可能 ※ただし、車両が走行することにより発生する背景の差分と区別するためには何らかの工夫が必要	①前のフレームで人を検出したフレームが誤検出 ②次のフレームで人を検出しなかったフレームが誤検出

物体検出 AI 冗長設計での「Trust AI Shield」モジュール案を表に示す。

表 物体検出 AI 冗長設計での「Trust AI Shield」モジュール案

No.	項目	説明	ユースケース
1	物体検出 AI(2 個)の AND をとる	異なる物体検出 AI を 2 つ利用し、両方の AI で共通して検出している結果のみ信頼できる検出として判断する(片方の AI でしか検出していない結果は出力しない)	①両方の AI が正しく検出する ②片方の AI は正しく、片方の AI は誤検出する ③両方の AI が誤検出する
2	物体検出 AI(2 個)の OR をとる	異なる物体検出 AI を 2 つ利用し、両方の AI で検出している結果をすべて信頼できる検出として判断する(片方の AI でしか検出していない結果も出力する)	①両方の AI が正しく検出する ②片方の AI は正しく、片方の AI は誤検出する ③両方の AI が誤検出する
3	物体検出 AI(3 個)の多数決	異なる物体検出 AI を 3 つ利用し、多数決で検出結果を判断する	①3 つ全ての AI が正しく検出する ②2 つの AI は正しく、1 つの AI は誤検出する ③1 つの AI は正しく、2 つの AI は誤検出する ④3 つ全ての AI が誤検出する

認識率の高い物体検出 AI での「Trust AI Shield」モジュール案を表に示す。

表 認識率の高い物体検出 AI での「Trust AI Shield」モジュール案

No.	項目	説明	ユースケース
1	環境条件を保証する	限定環境下での認識率が 99.9%であるので、その限定環境下であることを保証する。限定環境下であることを保証することで、結果的に AI が誤検出しないことを保証する。	-

制御量算出 AI での「Trust AI Shield」モジュール案を表に示す。

表 制御量算出 AI での「Trust AI Shield」モジュール

No.	項目	説明	ユースケース
1	物体検出センサと照合する	誤検出しない(機能安全等で保障されている等)物体検出センサの結果と照合し、矛盾が生じる場合は物体検出センサの結果を優先する。矛盾が生じない場合は AI の出力を信用する。	①AI と物体検出センサが両方正しく検出する ②AI は誤検出し、物体検出センサは正しく検出する

表に示したモジュール案の詳細をそれぞれ示す。

まず、物体検出 AI 単一設計パターンでの「Trust AI Shield」モジュール事例を示す。事例を示す上での前提条件として、対象システムは図のようなシステム構成を想定した。

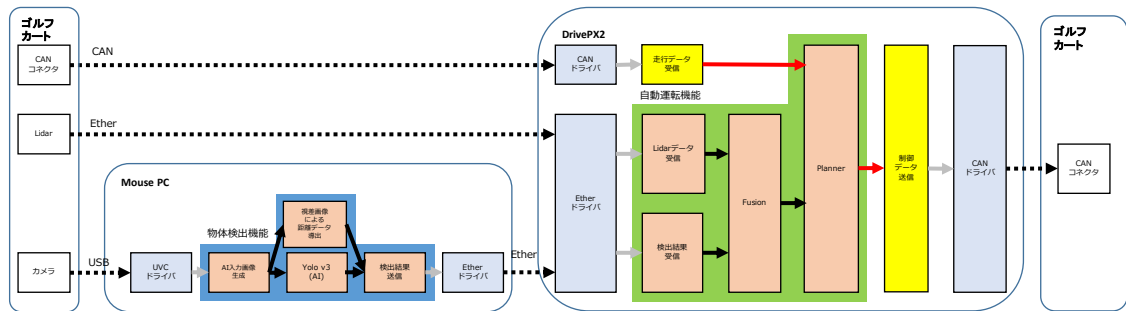


図 対象システム構成図

対象 AI の YOLO の概要を表に示す。

表 YOLO の概要

	項目	説明
名称	YOLO	darknet と呼ばれるニューラルネットフレームワークで実装された物体検出アルゴリズムおよび学習モデル
用途	物体検出	画像の中から定められた物体の位置と種別(クラス)を検出
入力	画像/ストリーム	画像もしくはストリーム画像を入力として利用可能
出力	物体の種別	検出した物体の種別(例：dog,car…)
	物体の位置	変換式を用いて、4 点の座標で取得可能
	信頼度	AI がどれくらい検出枠内に対象物体を含んでいるかおよび、検出枠が対象物体だとどのくらいの精度で予測したかを示す数値

本 AI を使用した場合に起こりうる誤検出のパターンは、以下 5 パターンを想定した。

- ・実際に物体が存在しないのに検出する
- ・実際に物体が存在するのに検出しない
- ・異なる物体として検出する
- ・検出箇所が実際の物体よりも大きい
- ・検出箇所が実際の物体より小さい

対象 AI システムの前提条件を表 3-3-2-2-7 に示す。

表 ACC システム前提条件

分類	前提条件
環境条件	視界が良好(雨、夜、霧等の視界不良がない)
	道路は常に平坦
	道路は急カーブおよび曲がり角がなく、見通しがよい
	片側車線のみ
	車線変更はなし(自車、他車共に)
	私道
	人および車両以外の障害物は存在しない
	人は「手に持ったマント等でカメラに対して写りこまないようにする」といった意図的にカメラに写りこまないような対応をしない
	人は同時に複数存在しない
性能条件	AI の物体検出処理速度は 60fps 程度
	自車両の走行速度は最大 20km/h(約 3m/sec)
センサ条件	センサは故障しないこと、もしくはセンサが故障したことはシステム側で検知可能であることが保証されている
	カメラの画角は 45°

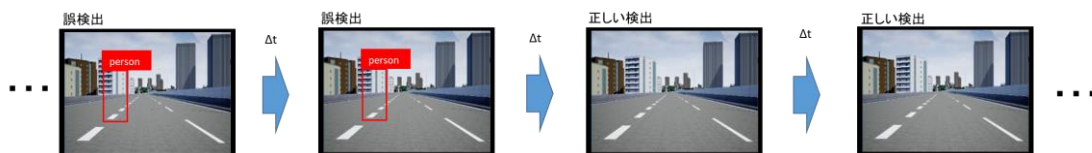
物体検出 AI 単一設計でのモジュール案の No. 1, 2 の詳細を示す。

①「画像真ん中付近で検出した人が次のフレームで検出しなくなる」

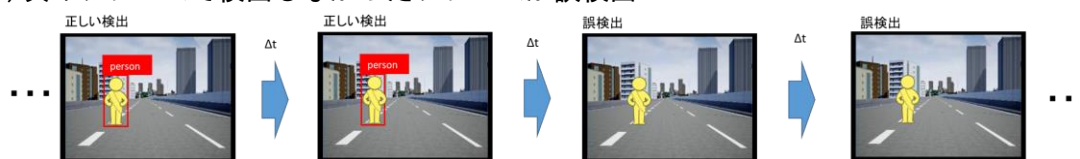
100m 走の世界記録は 9.58 秒であるため、1 フレーム間に人が移動可能な距離は、17cm である。画像真ん中付近で検出された人が 1 フレームでカメラ画角から外れることは不可能であるという考え方に基づいたモジュール案である。

考えられるユースケースを 2 パターン示す。

(1) 画像真ん中で人を検出したフレームが誤検出



(2) 次のフレームで検出しなかったフレームが誤検出

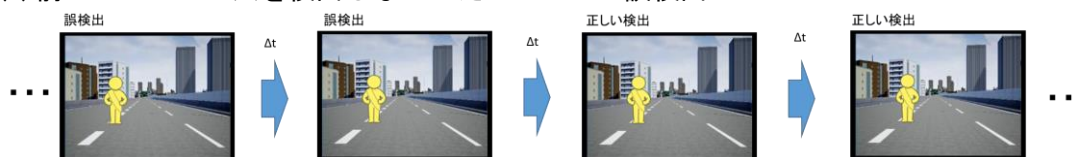


②「前のフレームで人を検出していなかったにも関わらず、次のフレームで人を検出する」

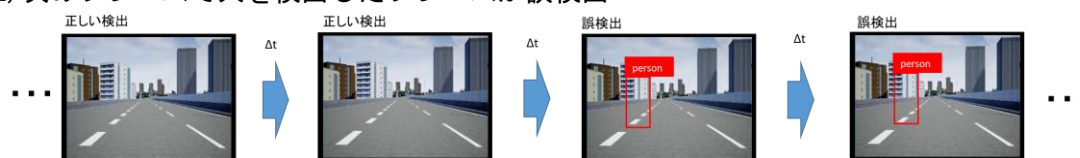
100m 走の世界記録は 9.58 秒であるため、1 フレーム間に人が移動可能な距離は、17cm である。画像真ん中付近で検出された人が 1 フレームでカメラ画角から外れることは不可能であるという考え方に基づいたモジュール案である。

考えられるユースケースを 2 パターン示す。

(1) 前のフレームで人を検出できなかったフレームが誤検出



(2) 次のフレームで人を検出したフレームが誤検出



次に、物体検出 AI 冗長設計パターンでの「Trust AI Shield」モジュール事例を示す。前提条件として、対象システムは図と同様のシステムを想定した。ただし、物体検出機能の AI モジュールは冗長設計であるとした。対象 AI には YOLO と、表に概要を示す SSD を想定した。物体検出 AI を 3 つ利用したパターンの事例も示すが、3 つ目の対象 AI は具体的なものは想定していないが、YOLO、SSD と同等の AI であると定義して検討を行った。

表 SSD の概要

	項目	説明
名称	SSD (SingleShotMultiBoxDetector)	畳み込みニューラルネットワーク(CNN:Convolutional Neural Network)を用いた物体検出アルゴリズム
用途	物体検出	画像の中から定められた物体の位置と種別(クラス)を検出
入力	画像 / ストリーム(※)	ソースコードを改変すればストリームでも対応可能
出力	物体の種別	検出した物体の種別(例: dog,car…)
	物体の位置	変換式を用いて、4 点の座標で取得可能
	信頼度	AI がどれくらい検出枠内に対象物体を含んでいるかおよび、検出枠が対象物体だとどのくらいの精度で予測したかを示す数値

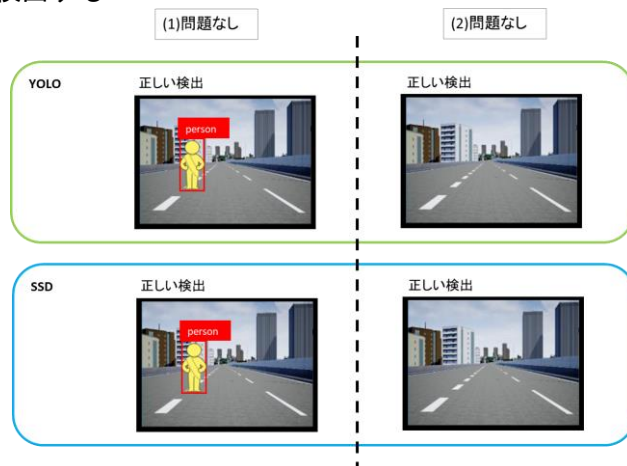
対象 AI システムの前提条件は表と同様である。

物体検出 AI 冗長設計でのモジュール案の No1～3 の詳細を示す。

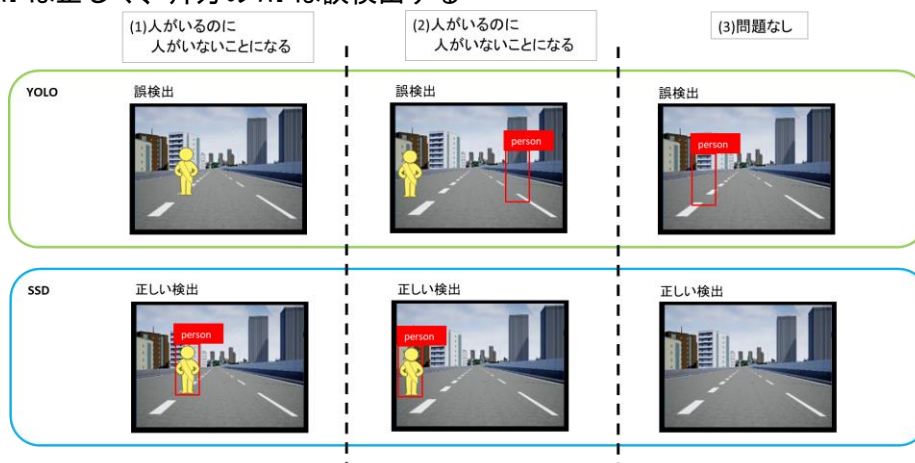
③「物体検出 AI (2 個) の AND をとる」

複数の AI を用いることにより冗長性を高めたモジュール案である。
考えられるユースケースを大きく 3 パターンに分けて示す。

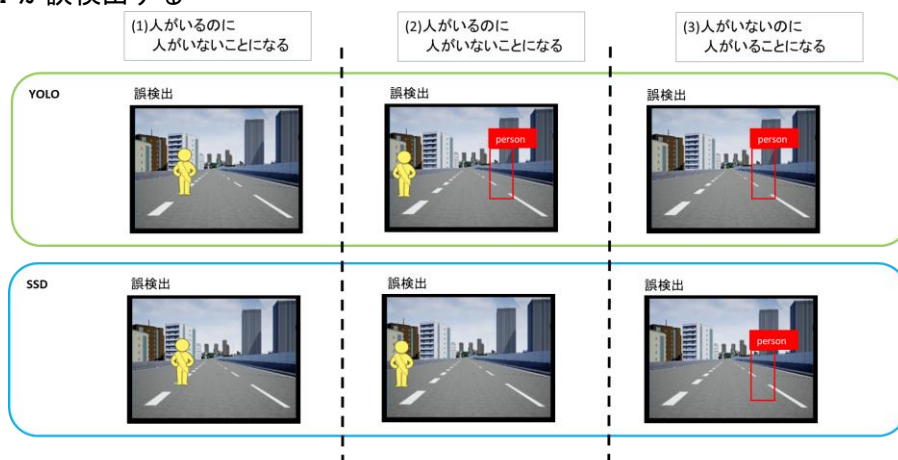
(1) 両方の AI が正しく検出する



(2) 片方の AI は正しく、片方の AI は誤検出する



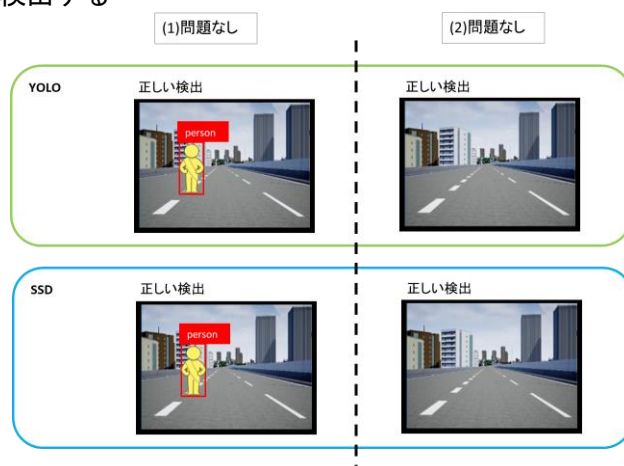
(3) 両方の AI が誤検出する



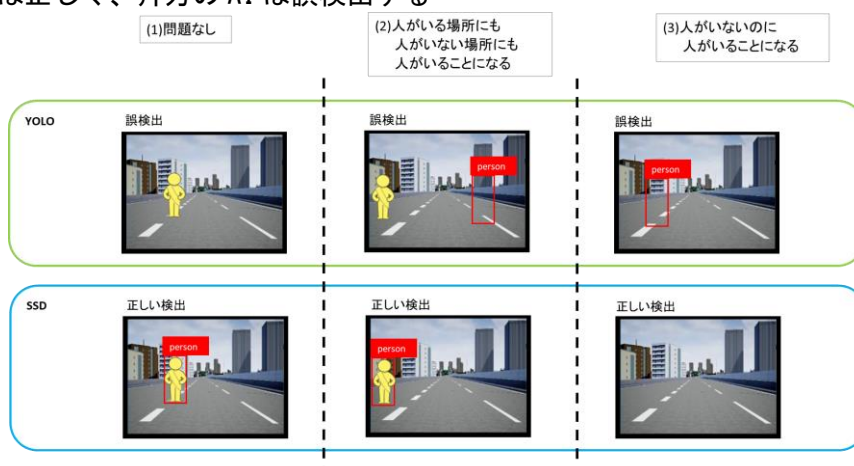
④「物体検出 AI (2 個) の OR をとる」

複数の AI を用いることにより冗長性を高めたモジュール案である。
考えられるユースケースを大きく 3 パターンに分けて示す。

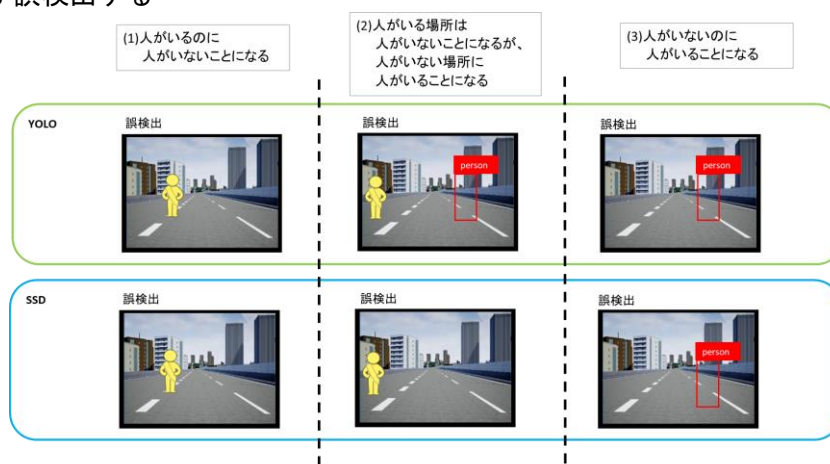
(1) 両方の AI が正しく検出する



(2) 片方の AI は正しく、片方の AI は誤検出する



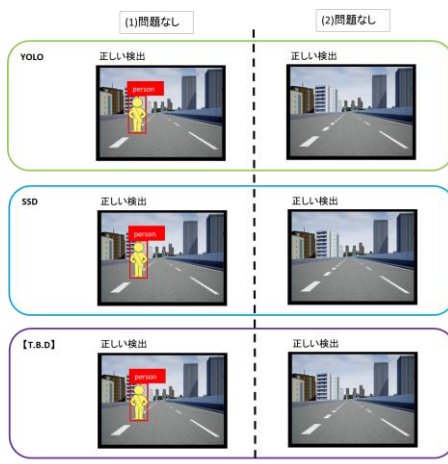
(3) 両方の AI が誤検出する



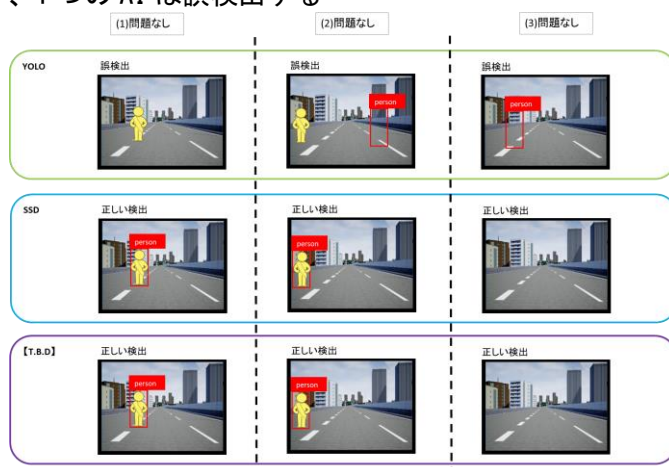
⑤ 「物体検出 AI (3 個) の多数決」

複数の AI を用いることにより冗長性を高めたモジュール案である。
考えられるユースケースを大きく 4 パターンに分けて示す。

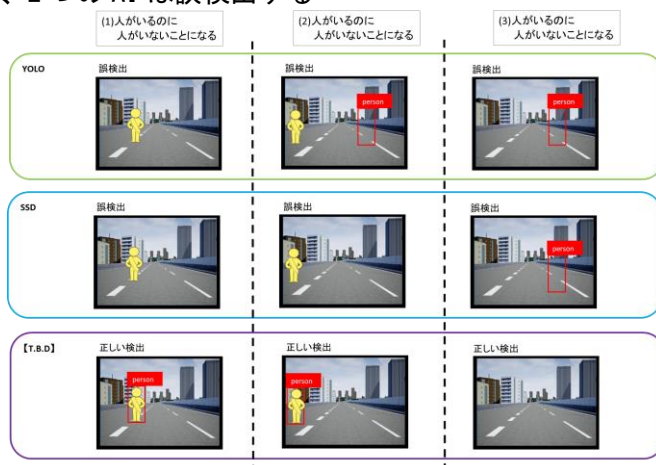
(1) 3 つ全ての AI が正しく検出する



(2) 2つのAIは正しく、1つのAIは誤検出する

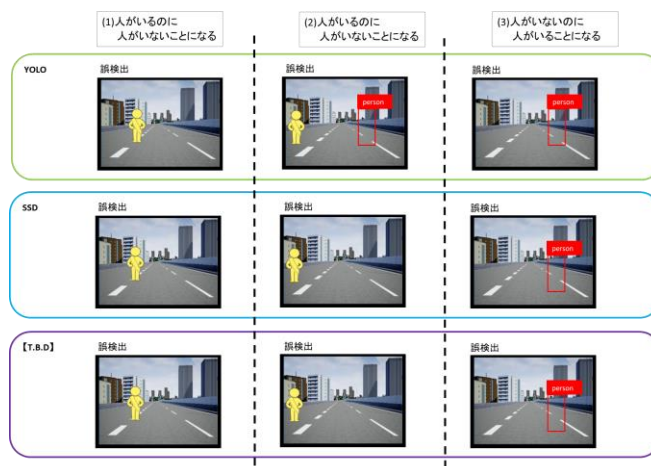


(3) 1つのAIは正しく、2つのAIは誤検出する



(4) 3つ全てのAIが誤検出する

（令和元年度戦略的基盤技術高度化支援事業）
「自律的自動運転の実現を支える人工知能搭載システムの安全性立証技術の研究開発」研究成果報告



認識率の高い物体検出 AI での「Trust AI Shield」モジュール事例を示す。前提として、対象システムは図 3-3-2-2-4 と同様のシステムを想定する。ただし、物体検出機能の AI モジュールには認識率 99.9%の物体検出 AI を使用すると仮定する。実際には本 AI は存在しないが、認識率以外の条件は表 3-3-2-2-7 と同等のものとする。

認識率の高い物体検出 AI でのモジュール案の詳細を示す。

①「環境条件を保証する」

環境条件下では認識率 99.9%であるので、限定環境下であることを TAS モジュールで保証するモジュール案である。

最後に制御量算出 AI での「Trust AI Shield」モジュール事例を示す。前提として、基本的なシステム構成は、「廃線後を利用した無人車両による限定区間内ピストン輸送システム」を想定する。ACC 機能および LKAS 機能部分に、制御量算出 AI を用いたシステムを想定する。

対象 AI である Donkey についての概要は、表に示す。

表 Donkey の概要

	項目	説明
名称	Donkey	TensorFlow を使用したニューラルネットワークの AI
用途	車両制御	手動走行時のデータを教師データとして学習し、カメラ画像からステアリングとスロットルの値を出力することで RC カー等を走行させる
入力	画像	ソースコードを改変すればストリームでも対応可能
出力	スロットル	DonkeyCar のスロットル
	ステアリング	DonkeyCar のステアリング

本 AI を使用した場合に起こりうる誤検出は、以下 3 パターンを想定する。

- ・ 専有道路を逸脱するステアリングとスロットルを出力する
- ・ 障害物に衝突するステアリングとスロットルを出力する
- ・ 走行できる状態で走行せずにスロットルが出力されない

対象 AI システムの前提条件は表に示す。

表 MaaS システム前提条件

分類	前提条件
環境条件	1 本道でカーブはあるが右左折を要する曲がり角は無し
	カーブの曲率半径は R=100 以上 ※R=100 はカーブ半径が 100m
	勾配は 0 度～2 度(3%未満) ※鉄道の基準において勾配対策が不要とされる範囲
	専有路内は一定区間毎にカメラを設置し専有路を監視
	専有路と区間外の仕切りは簡易の柵など、人・動物の立ち入りが不可能ではない程度の物理的閉鎖
	一般道との交差は無し
	専有路への一般車両の立ち入りは無し
	専有路への動物の立ち入りはあり
	専有路への人の立ち入りは原則禁止だが、緊急時は運用員の立ち入りがあり(障害物除去など)
	障害物はある(立ち入った動物など)
性能条件	AI の処理速度は 60fps 程度

制御量算出 AI でのモジュール案の詳細を示す。

①「物体検出センサと照合する」

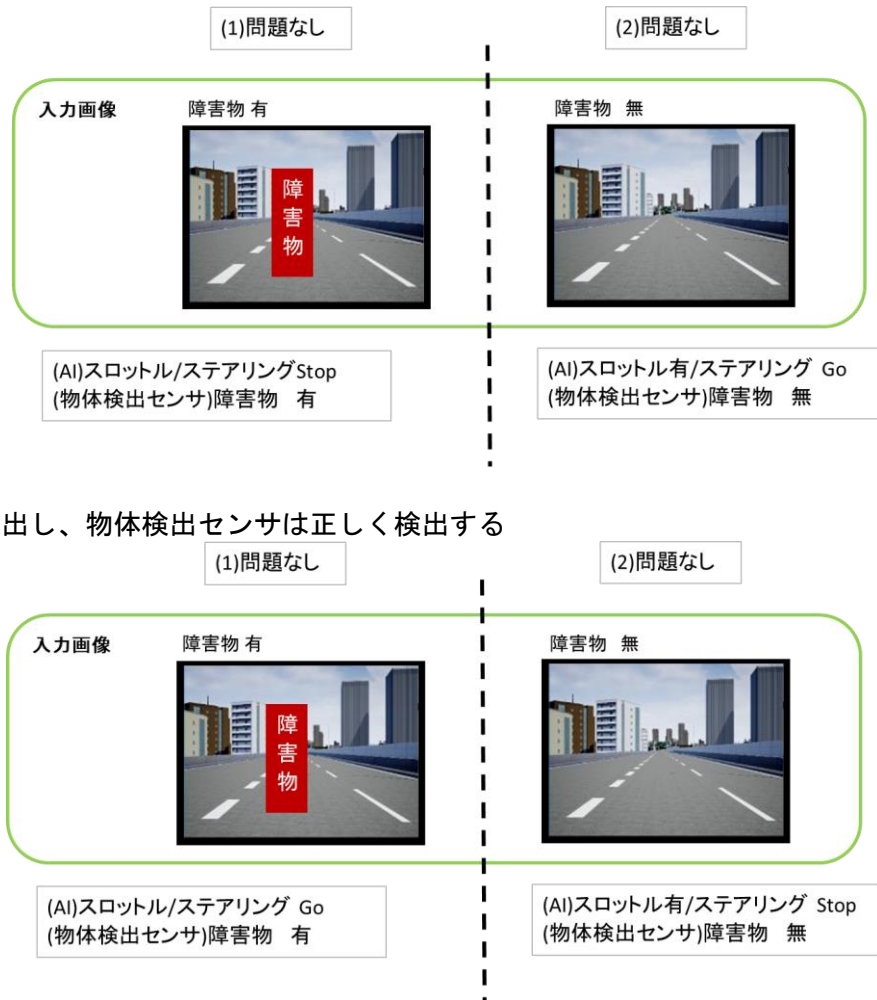
物体検出センサの結果を照合することで、“AI が障害物に衝突するステアリングとスロットルを出力したこと”を検出する。機能安全などの別のプロセスで出力が保証されている物体検出センサを用いることで、出力は保証されるという考えに基づいたモジュール案である。加えて本 AI は連続値を扱う結果のため、「Trust AI Shield」モジュールと AI の出力の間で矛盾が生じない場合は、AI のメリットを生かすことが可能である。

物体検出センサが誤検出することは想定せず、考えられるユースケースを 2 パターン示す。

なお、DonkeyAI の出力は本来連続値であり 2 値では表現できない。ユースケースを示す場合には、以下の定義に基づいて示す。

- ・そのまま進む場合をステアリング/スロットル : Go
- ・停車またはステアリングによる回避の場合をステアリング/スロットル : Stop

(1) AI と物体検出センサが両方正しく検出する



4.3.3. 「Fuzzing for AI」

4.3.3.1. 「Fuzzing for AI」 設計

「Fuzzing for AI」の目的は以下の2つである。

- ・ AI を事前に解析することで AI 搭載システムの機能不全を未然に防ぐ
- ・ 効率的に現実的なコストでの検証を実施すること

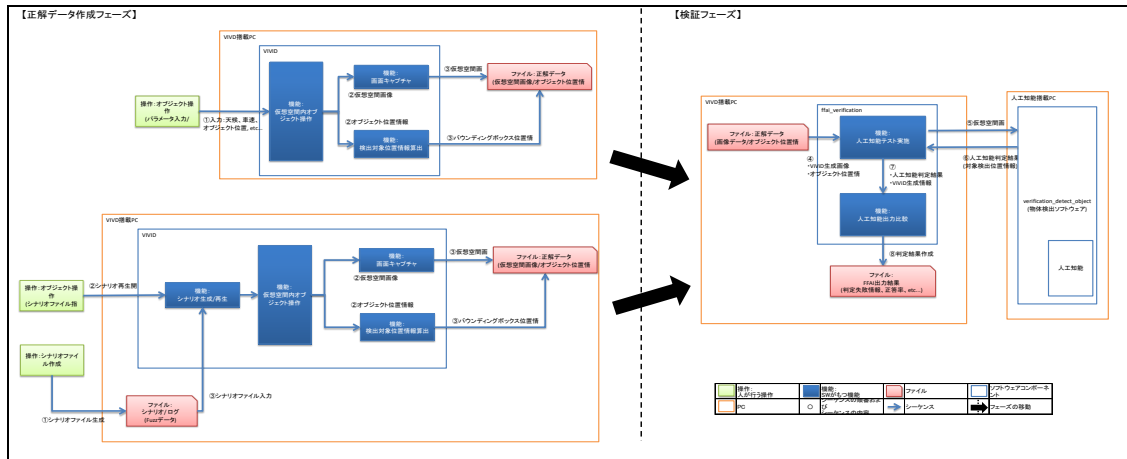
ViViD を利用することで正解付きデータを取得することで、多くのテストケースを自動で検証できるため、コストダウンが望める。さらにオプションとして、ノイズ付与機能を用意している。これによりテストケースのバリエーションを増やすことができるとともに、ノイズ付与前後の結果を比較することで、ノイズへの耐性に関するリスク分析を可能にする。本研究では実証実験システムに利用した物体検出 AI を対象とした、効率的に現実的なコストで検証できるツールの開発を完了した。

「Fuzzing for AI」は、大きく分けて正解データ作成フェーズと検証フェーズの2つに分けることができる。「Fuzzing for AI」の構成を図 3-3-3-1-1 に示す。正解データ作成フェーズは仮想空間内オブジェクトの操作方法が異なる、下記の2パターン存在する。

- ・ 手動操作
- ・ 事前に作成したシナリオファイルで操作

手動操作では即座にデータを取得することができる。一方でシナリオファイルを利用すると、事前にシナリオファイルの用意が必要になるが、より細かなオブジェクトの動きが指定できるだけでなく、同一のシナリオファイルを使用することで再現することも可能であるといったメリットがある。

(令和元年度戦略的基盤技術高度化支援事業)
「自律的自動運転の実現を支える人工知能搭載システムの安全性立証技術の研究開発」研究成果報告



図「Fuzzing for AI」構成図

物体検出 AI で検出する場合、検出位置に誤差が生じるという特性がある。検出枠と正解オブジェクトの比較の際には、補正を行うためのマージンを設定した。マージン補正のイメージを図に示す。マージンの許容値は、検証対象の AI およびシステムに依存するため、マージンを検証時に設定可能とする。

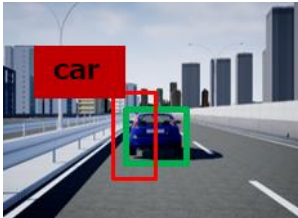
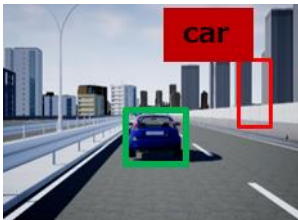
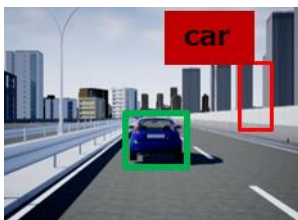


図 マージン補正イメージ

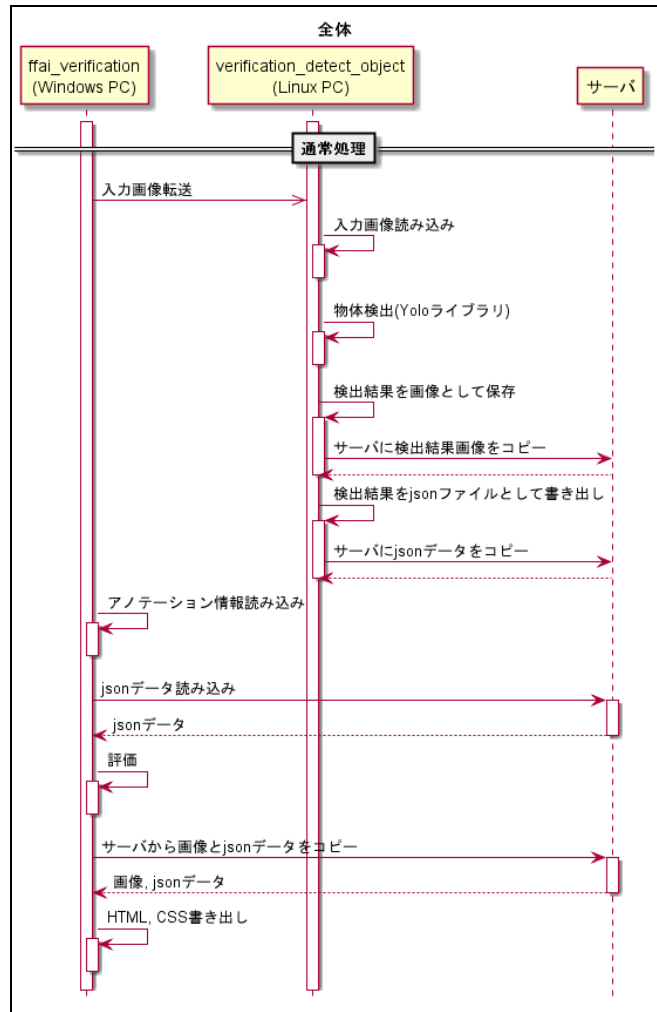
評価に用いる評価項目のパターンを表に示す。表の例(4)と(6)は同一画像であり、「検出位置異常」と「認識種別異常」の両方が発生している。このように、1つの検出枠において評価項目が重複して該当する場合も想定する。例では1枚の画像に対して1つの検出枠しか記載していないが、1枚の画像に対して複数の検出枠がある場合も想定する。

表「Fuzzing for AI」評価項目

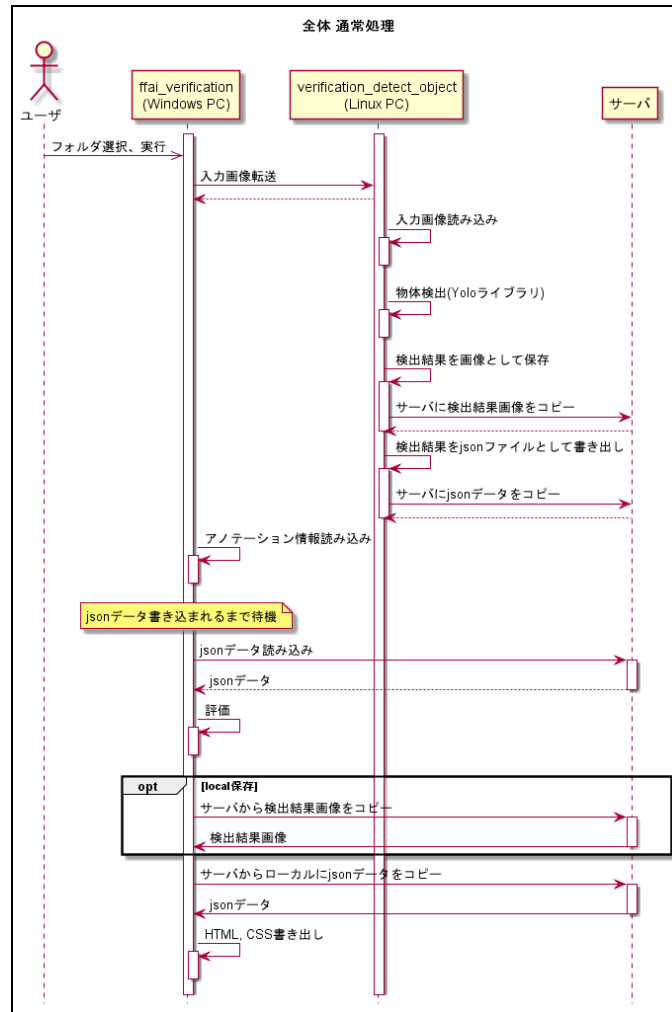
評価項目	定義	例
検出サイズ大	物体検出結果の座標すべてが、正解オブジェクトの座標（マージン上下左右付加後）より外側	凡例 緑色の枠 ：正解オブジェクト（マージン上下左右付加） 赤色の枠 ：物体検出結果 (1)

検出サイズ小	物体検出結果の座標すべてが、正解オブジェクトの座標(マージン上下左右付加後)より内側	(2) 
検出位置異常	物体検出結果の座標の一部が、正解オブジェクトの座標(マージン上下左右付加後)より外側または内側	(3)  (4) 
認識種別異常	物体検出結果の座標が、正解オブジェクトの座標(マージン上下左右付加後)を少なくとも一部包含している状態かつ、認識の種別が異なる	(5)  (6) 
未検出	正解オブジェクトの座標(マージン上下左右付加後)が、物体検出結果の座標に全く包含されずに存在する	(7)  (8) 
誤検出	物体検出結果の座標が、正解オブジェクトの座標(マージン上下左右付加後)を、全く包含せずに存在する	(9)  (10) 

「Fuzzing for AI」の全体シーケンス図を図 (a) に示す。図 (a) に記載されている通常処理部分は図 (b) に示す。

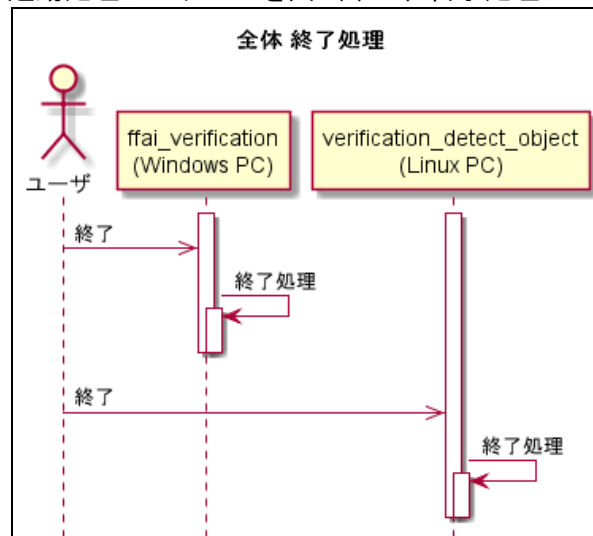


図(a) 「Fuzzing for AI」全体シーケンス

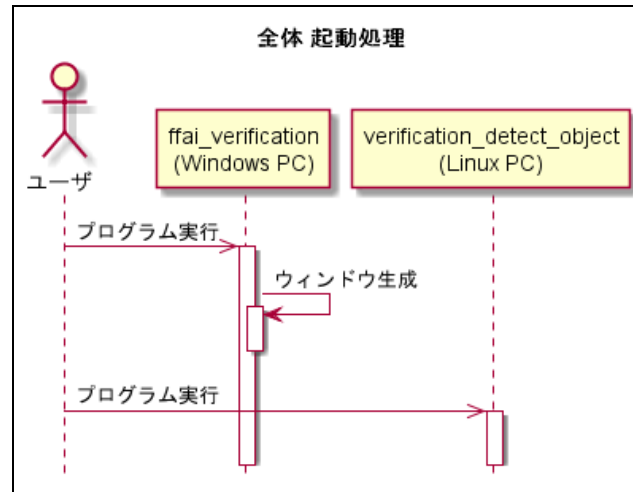


図(b) 「Fuzzing for AI」 全体シーケンス 通常処理

「Fuzzing for AI」の起動処理シーケンスを図 (a)に、終了処理シーケンスを図 (b)に示す。

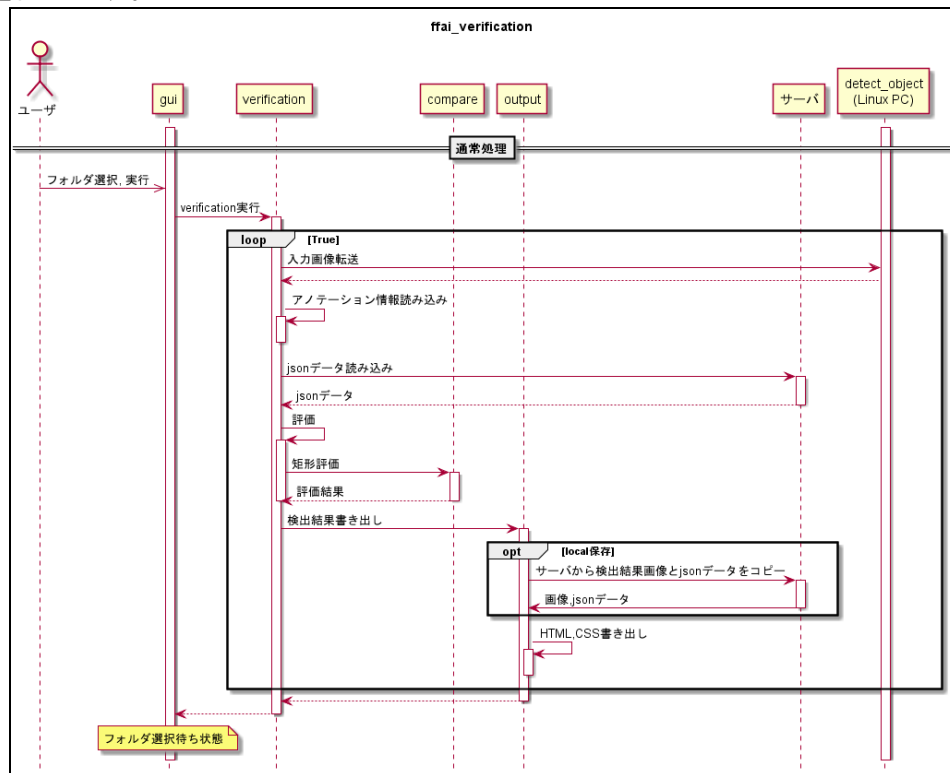


図(a) 「Fuzzing for AI」 起動・終了シーケンス起動処理



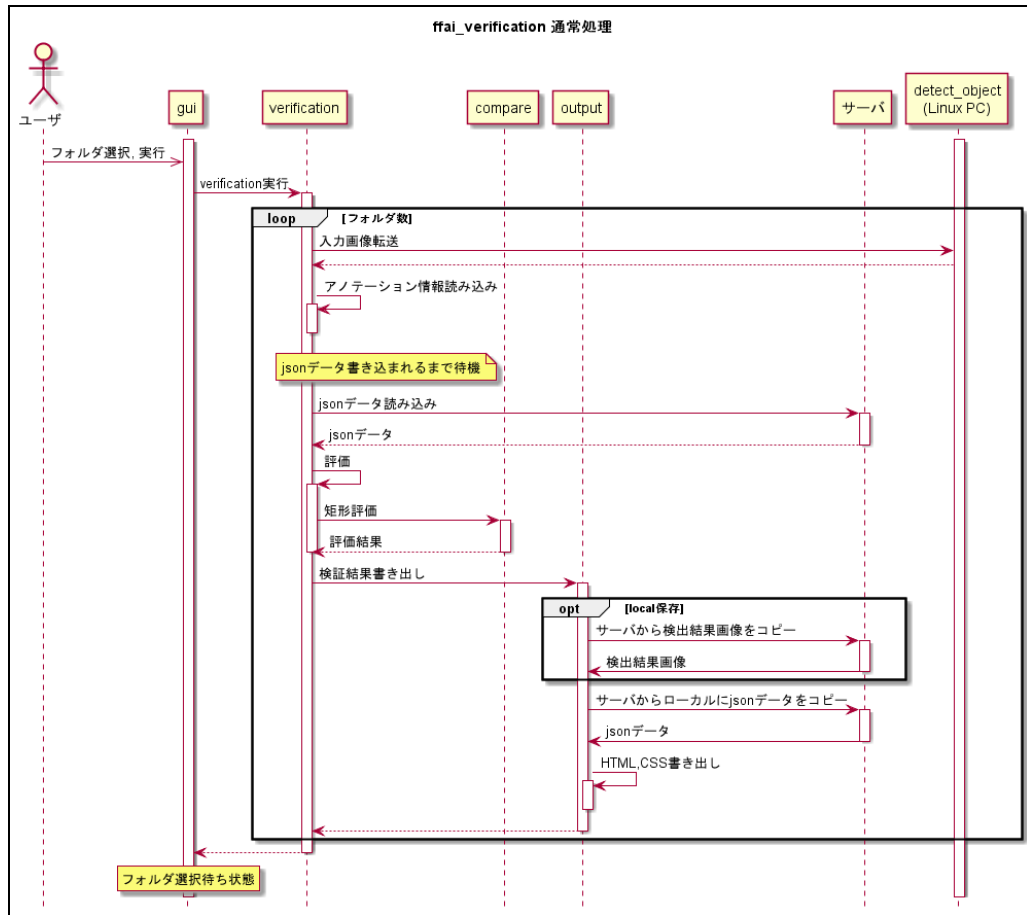
図(b) 「Fuzzing for AI」 起動・終了シーケンス終了処理

「Fuzzing for AI」の実行および検証を実施するアプリケーション「ffai_verification」の全体シーケンスを図に示す。



図「ffai_verification」全体シーケンス

通常処理部分を図に示す。



図「ffai_verification」通常処理シーケンス

「Fuzzing for AI」の起動処理シーケンスを図 (a)に、終了処理シーケンスを図 7(b)に示す。

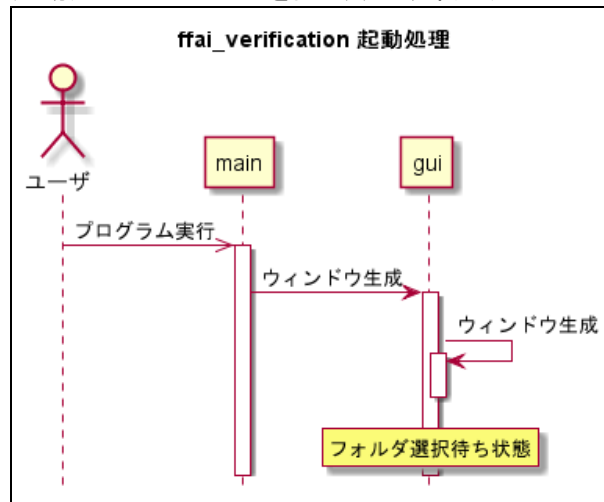
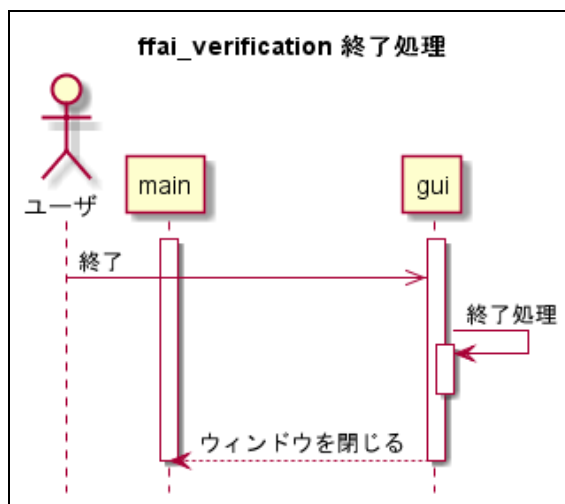
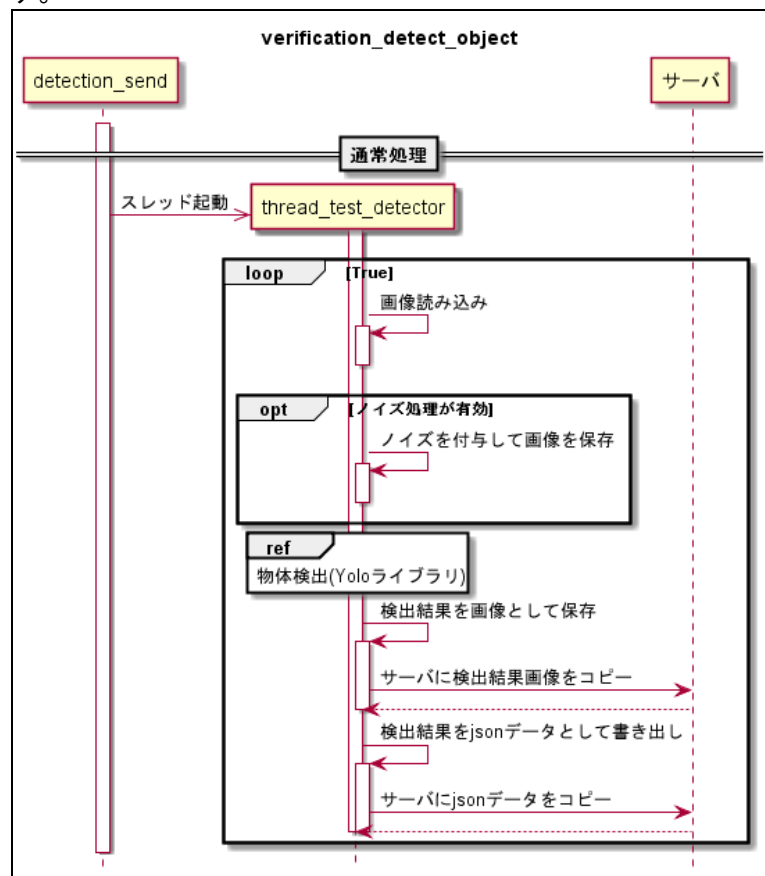


図 (a) 「ffai_verification」起動・終了シーケンス 起動処理

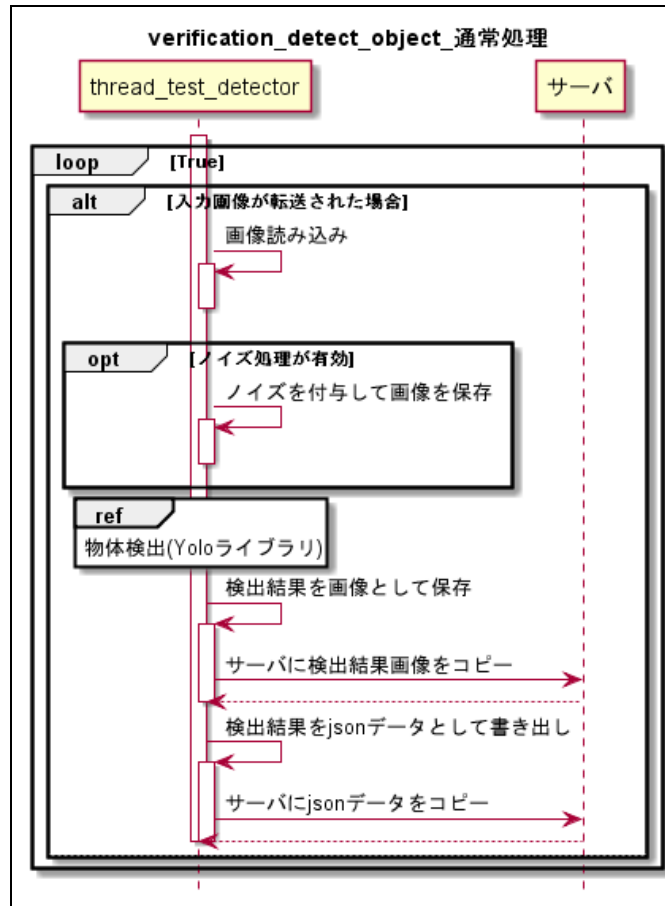


図(b) 「ffai_verification」 起動・終了シーケンス 終了処理

次に、「Fuzzing for AI」の物体検出を実行し、検出結果を格納するアプリケーション「verification_detect_object」の全体シーケンスを図 (a)に示す。図 (a)に記載されている通常処理部分は図 (b)に示す。

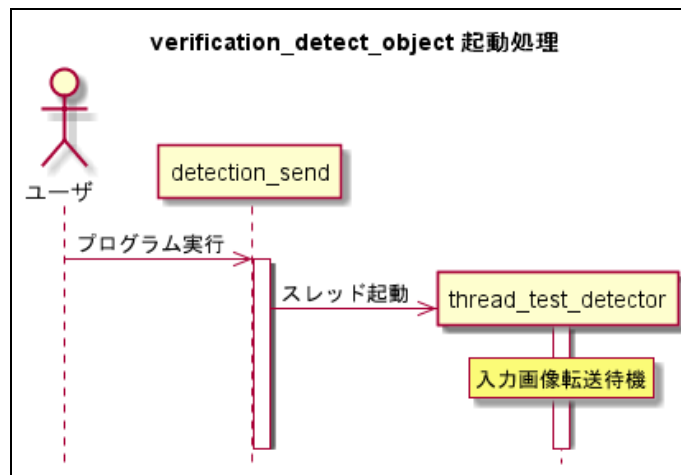


図(a) 「verification_detect_object」 全体シーケンス



図(b) 「verification_detect_object」全体シーケンス 通常処理

「verification_detect_object」の起動処理シーケンスを図 (a) に、終了処理シーケンスを図 3(b) に示す。



図(a) 「verification_detect_object」 起動・終了シーケンス

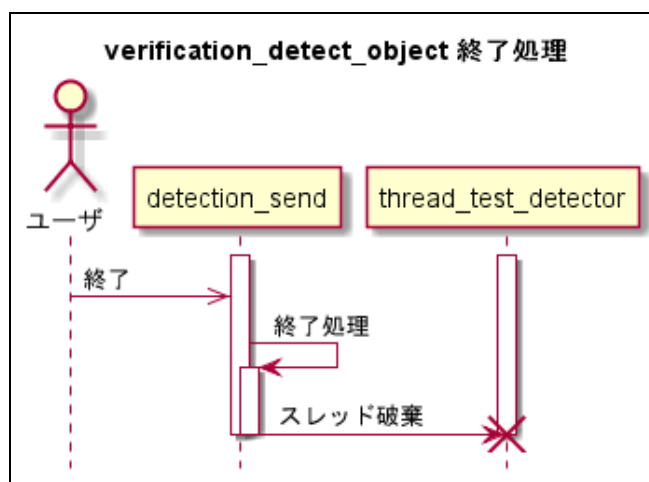


図 (b) 「verification_detect_object」 起動・終了シーケンス

4.3.3.2. 「Fuzzing for AI」 使用方法

ViViD を利用して作成した正解付きデータの 1 つを図に示す。図 (a) の画面キャプチャと、図 (b) のアノテーション情報が記載された xml ファイルが対になって出力される。画面キャプチャは物体検出 AI への入力として使用し、アノテーション情報は「Fuzzing for AI」の検証時の正解データとして使用する。

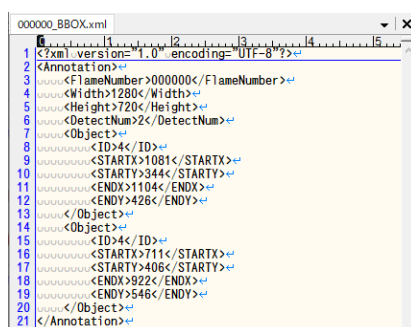
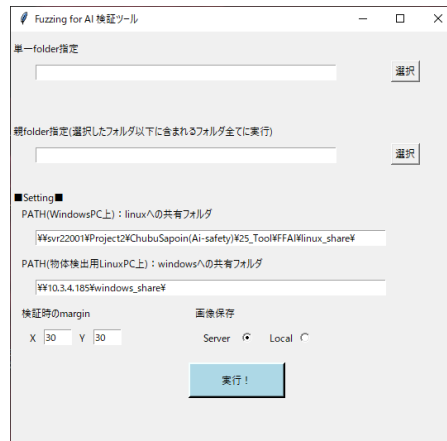


図 (a) 正解付きデータ 画面キャプチャ



図 (b) 正解付きデータ アノテーション情報

「ffai_verification」の実行画面を図に示す。図の情報をまとめて格納したフォルダを指定することで、フォルダ以下に格納されているデータすべてを対象として検証を実行する。



図「ffai_verification」実行画面

4.3.3.3. 「Fuzzing for AI」検証結果

「ffai_verification」のテスト実施結果の一部を図群に示す。「Fuzzing for AI」は検証結果をhtml形式と、json形式の2通りで出力する。

図は1回のテスト実施時に検証した結果を項目ごとに集計した結果である。

項目	Frame									
Folder	全フレーム数	検出に成功したフレーム数	検出に失敗したフレーム数	検出サイズ大フレーム数	検出サイズ小フレーム数	検出位置異常フレーム数	認識種別異常フレーム数	未検出フレーム数	誤検出フレーム数	検出数相違フレーム数
2019-12-10-16-47-44	1051	657	394	0	0	79	29	319	0	347

図「ffai_verification」検証結果の summary

図は物体検出 AI に入力した画像ごとの結果をまとめた結果である

項目	Result								
Frame	0:失敗,1:成功	検出サイズ大数	検出サイズ小数	検出位置異常数	認識種別異常数	未検出数	誤検出数	検出数相違	出力画像
000628_RGB	0	0	0	0	0	1	0	1	Image
000629_RGB	1	0	0	0	0	0	0	0	Image
000630_RGB	0	0	0	0	0	1	0	1	Image
000631_RGB	0	0	0	0	0	1	0	1	Image
000632_RGB	1	0	0	0	0	0	0	0	Image
000633_RGB	0	0	0	1	0	0	0	0	Image

図 3-3-3-2 「ffai_verification」 frame ごとの検証結果

図は正解アノテーション情報のオブジェクトごとにまとめた結果である。評価項目がすべて OK の場合はセルが緑になるが、評価が1つ以上 NG の場合はセルが赤かつ問題の項目セルが黄色になる。

項目		Result						
Frame	Object	0:失敗,1:成功	検出成功	検出サイズ大	検出サイズ小	検出位置異常	認識種別異常	未検出
000628_RGB	0	1	1	0	0	0	0	0
000628_RGB	1	0	0	0	0	0	0	1
000629_RGB	0	1	1	0	0	0	0	0
000629_RGB	1	1	1	0	0	0	0	0
000630_RGB	0	1	1	0	0	0	0	0
000630_RGB	1	0	0	0	0	0	0	1
000631_RGB	0	1	1	0	0	0	0	0
000631_RGB	1	0	0	0	0	0	0	1
000632_RGB	0	1	1	0	0	0	0	0
000632_RGB	1	1	1	0	0	0	0	0
000633_RGB	0	1	1	0	0	0	0	0
000633_RGB	1	0	0	0	0	1	0	0

図「ffai_verification」 object ごとの検証結果

検出成功の例として「000628_RGB」の出力画像を図(a)に、検出失敗の例として「000629_RGB」の出力画像を図(b)に示す。図(a)では車両および自転車の2つをマージン内で検出しているので検出成功になるが、図(b)では自転車を検出できなかったためこのフレームでは検出失敗となった。



図(a)「verification_detect_object」出力画像 検出成功



図(b)「verification_detect_object」出力画像 検出失敗

同様のテストを手作業で実施した場合との比較を、表に示す。手作業で検証を行うと1時間当たり約35件で実施できるところを、ツールを使用すると1時間当たり約3600件の検証ができた。この結果から「Fuzzing for AI」を使用することで、効率的に現実的なコストでの検証を実現できたといえる。

FFAI 検証ツールは手作業とは異なり、個人の能力に依存しない制度で検証可能である。

表「Fuzzing for AI」検証結果比較

	「Fuzzing for AI」 検証ツール	手作業
正解付き画像データ生成	共通	
AIに画像データを入力し、結果を取得	時間：約 3,617 件/h 精度：100%	未実施
AIの出力と正解データの比較		時間：約 35 件/h 精度：個人の能力に依存

5. 補助事業の成果に係る事業化展開について

5.1. 想定している具体的なユーザー、マーケット及び市場規模等に対する効果

我が国の首相が「2020 年までに運転手が乗車しない自動車によって地域の人手不足や移動弱者を解消すべき」と発言され、高速道路における無人トラック走行や無人バス・タクシーなどの実証実験を実施するとの方針が示されたように、当該産業分野での自律的自動運転の開発技術は非常に重要であり、かつ他国に先駆けて実現しなければならない喫緊の状況にある。

No	カテゴリ	概要	市場規模（億円）		
			2015 年	2020 年	2030 年
1	農林水産業関係		28	316	3,842
		農林水産業用ロボティクス市場等	28	316	3,842
2	製造業関係		1,129	29,658	121,752
		産業用ロボティクス市場等	60	6,164	17,571
		自動運転車製造市場	1,069	23,494	104,181
9	運輸業関係		1	46,075	304,897
		オンデマンド・モビリティ市場	0	8,630	106,449
		自動運転トラック輸送市場	1	37,445	198,448

引用：EY 総研 人工知能が経営にもたらす「創造」と「破壊」より

日本における人工知能の全体的な市場規模は 2015 年には 3.7 兆円であったが、2020 年には 23.6 兆円、2030 年には 86.9 兆円と予測されており、とりわけ、自律的自動運転技術による、旅客および運輸分野への計上が多く、人工知能への期待が寄せられている。

5.2. 事業化見込み（目標となる時期・売上規模）

製品等の名称		人工知能搭載システムの安全対応コンサルティング事業				
開発事業者		株式会社ヴィッツ、株式会社アトリエ				
想定するサンプル出荷先		トヨタ自動車株式会社、アイシン精機株式会社、ヤマハ発動機株式会社				
スケジュール	事業年度	平成 32 年度	平成 33 年度	平成 34 年度	平成 35 年度	平成 36 年度
	サンプルの出荷・評価	→				
	追加研究	→	→	→		
	設備投資					
	製品等の生産					
	製品等の販売					
	特許出願	→				
	出願公開				→	→
	特許権設定				→	→
	ライセンス付与					
売上見込	売上高（千円）	10,000	20,000	100,000	200,000	300,000
	販売数量	2 案件	4 案件	10 案件	20 案件	30 案件

（令和元年度戦略的基盤技術高度化支援事業）
「自律的自動運転の実現を支える人工知能搭載システムの安全性立証技術の研究開発」研究成果報告

	売上高の根拠	現在株式会社ヴィッツが行っている、機能安全にかかわるコンサルティング事業は1件あたり500万円～1000万円で受託している。 平成34年度に受託件数が増える理由としては、平成36年度、西暦2024年に各自動車メーカーがレベル4自動運転を実現することを計画しているため、2年前にあたる平成34年度に受託件数が増加することを想定している。
--	--------	---

製品等の名称		人工知能搭載システムに対応した安全機能ソフトウェア部品販売事業（「Trust AI Shield」および「Fuzzing for AI」）				
開発事業者		株式会社ヴィッツ、アーク・システム・ソリューションズ株式会社				
想定するサンプル出荷先		トヨタ自動車株式会社、アイシン精機株式会社、ヤマハ発動機株式会社				
スケジュール	事業年度	平成32年度	平成33年度	平成34年度	平成35年度	平成36年度
	サンプルの出荷・評価	→				
	追加研究	→	→	→		
	設備投資					
	製品等の生産	→	→	→	→	→
	製品等の販売	→	→	→	→	→
	特許出願	→				
	出願公開				→	→
	特許権設定				→	→
	ライセンス付与					
売上見込	売上高（千円）	10,000	20,000	100,000	150,000	200,000
	販売数量	10本	20本	20本	20本	20本
	売上高の根拠	現在株式会社ヴィッツが販売している機能安全対応の保護機能ライブライはおおよそ1本あたり100万円で販売を実施している。またセキュリティ技術で使われるFuzzing Test ツールも同じく1本あたり100万円の価格を設定している。 平成34年度に販売数量が変化せず売上高が増える理由としては、平成36年度、西暦2024年に各自動車メーカーがレベル4自動運転を実現することを計画しているため、人工知能搭載システムの安全性立証と関連したソフトウェア開発業務の付帯が増える見込みから増額している。				

5.3. 事業化に至るまでの遂行方法や今後のスケジュール

本研究成果を活用した事業化は大きく2事業となる。

一つは、人工知能搭載システムの安全性立証技術に伴う知財提供、いわゆる、コンサルティング事業である。もう一つは、人工知能搭載システムの安全対策を実現する安全機能ソフトウェア部品の販売である。

上記二つの事業を実現するためには、国際認証機関（欧州もしくは米国）との技術ディスカッション、前述機関からのコンサルテーション提供による知識の集積などを確実に実施し、2年度内に国内企業向けのコンサルテーション事業を開始しなければ2025年前後の製品化に対応できないため、素早い対応が必要となる。

また、安全機能ソフトウェア部品の販売は、魅力的なソフトウェア部品であるため、これらのソフトウェア部品を利用するために必要となる“インテグレーション”が本研究事業企業には大きな収益を生む事業とだと考えられる。当面の対象市場は自律的自動運転を実現する自動車関連メーカーへの知財・ソフトウェア部品提供市場となる。販売および事業化戦略としては、戦略的基盤技術高度化支援事業による研究開発であることを最大限アピールする。

5.4. 成果（試作品）の無償譲渡や無償貸与

川下企業である国内自動車メーカーおよび国内自動車部品供給メーカーは、来るべき2020年東京オリンピックでの自律的自動運転技術の全世界へのアピールを実現するために、自律的自動運転技術の高度化・品質（アピールできるレベルの限られた環境下での確実性と安全性）を実現するための活動で手一杯となり、日本国内の四季環境や製品に向けた各種規格適応などの対策には手が付けられない状況にある。

そのため、自動運転の本流である自律的自動運転技術の高度化はメーカー各社が責任を負うべき技術領域であり、規格適応、特に安全に関する規格への適応などは対策方法をコンサルテーション、関連技術文書の購入等により導入したいと考えている。すなわち、メーカー自身が蓄積すべき技術は自律的自動運転技術の高度化であり、製品に必要な規格適応、とりわけ対応が困難な安全技術に関する規格適応方法は他社からの導入を希望している。

我々の成果を確実な利用につなげるため、研究成果の前出しを積極的に検討する。

6. 補助事業の成果に係る知的財産権等について

6.1. 知的財産権の出願及び取得並びに論文掲載の有無

令和元年度研究活動においては、知的財産に関連する出願及び取得は実施していない。また、論文への掲載なども実施していない。

6.2. ライセンス契約等による事業展開

令和元年度研究活動においては、ライセンス契約等による事業展開は実施できていない。