優秀賞

D4 チームは自動運転向け画像認識システムを対象として AI システム特有の攻撃手法に対する対策案を検討しました。
具体的には自動運転における脅威のリスク分析と STAMP/STPA による安全性分析を行い，画像認識システムに対する攻撃手法と
その対策を洗い出し、実験的調査でセキュリティ要件を策定し防御手法の有効性を確認しました。

画像認識システムに限らず AI システムは産業応用で広く使われるようになってきており、安心・安全に利用できることが求められてきています。
AI システムにおけるセキュリティの脅威は従来型に加えて AI 特有の脆弱性を突いた攻撃に対処しなければなりません。
こうした AI システムに対するセキュリティガイドラインは存在しているが、対策については十分に示されていません。
本チームでは自動運転車向け ML 画像認識を対象として CRSS を用いて脅威の優先度を示し、STAMP/STPA によりハザードとその要因と対策を整理したこと。画像認識で利用される Yolo に対して攻撃を実際に検証して、危険性を確認し、防御に対する検討をしたことが実用上有効と考えられる初期の知見を得ました。
今後の発展すると想定される課題に対して先進的な取組である点が高く評価されました。

以上を高く評価して優秀賞を与えることとします。

国立情報学研究所 GRACE センター長・特任教授

本位田真一