# Reliable Authentication of Neural Networks

*Reliable algorithm to determine adversary example free zones for artificial neural networks*

CHALLENGES: Secure societies – Protecting freedom and security of Europe and its citizens

## PROBLEM DESCRIPTION

Automatic licence plate recognition systems are based on artificial neural networks. Surprizingly small data perturbations, the so-called adversarial examples can result in false classifitation. This research aims to have a reliable verification algorithm.

## CHALLENGES AND GOALS

We have more and more artificial neural network based automated recognition systems. In recent years more and more problems have been reported: adversarial examples can be given and the well trained NN systems give bad classification.

We aimed to give an interval arithmetic based reliable algorithm that is capable to determine large adversary example free

Zones. We planned to check the limitation of interval arithmetic based reliable verification algorithm

PRODUCTIVE SECTOR: Transportation industry, Healthcare systems

## MATHEMATICAL AND COMPUTATIONAL METHODS

- Basic technique: check the classification relations by an interval arithmetic based algorithm
- Input: a well trained artificial network and an image of known correct classification
- Output: a large set of images that have some additive noise and all of them are adversarial example free, i.e. they are correctly classfied

*Practical roblem field: licence plate recognition*

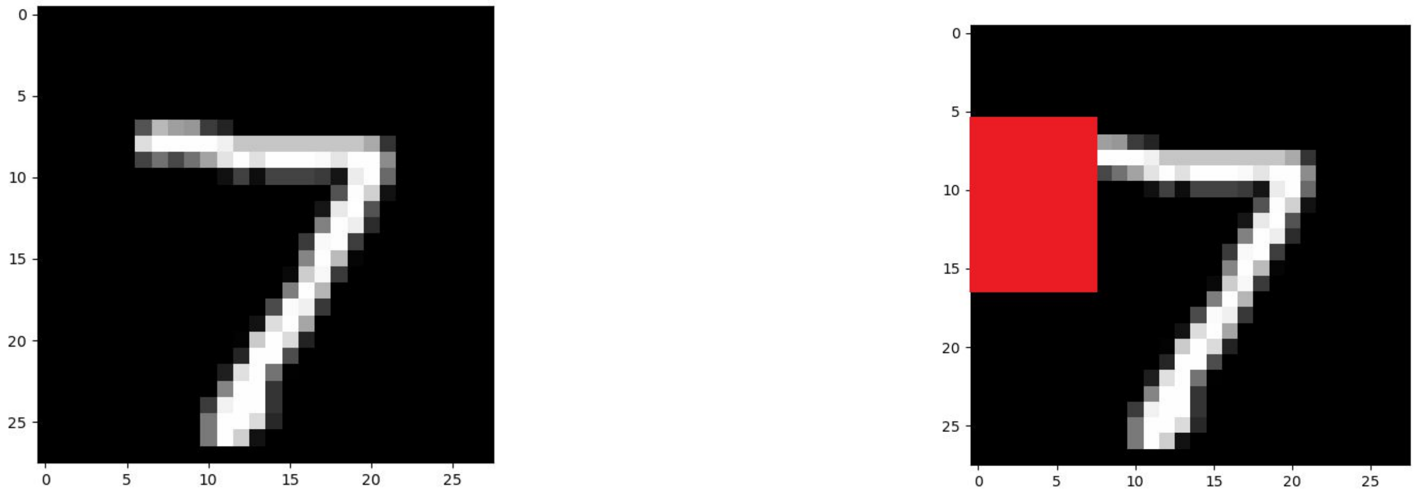
# Securing Cloud Authentication

*Design and analysis of a distributed authentication protocol for cloud services*

## Results and Benefits

We have designed, implemented and tested and interval arithmetic based reliable verification algorithm that is capable to check simple artificial neural networks and realistic size images.

The involved firm can improve its licence plate recognition application , and provide more reliable solutions for the costumers.



Arbitrary values can be in the red rectangle without destroying the correct classification

By applying interval arithmetic, we can determine the level of noise applied on images that will not change a correct classification of a given trained artificial neural network

Redink Ltd.

SZTE TTIK INFORMATIKAI INTÉZET

Institute of Informatics, University of Szeged