

Verificarea rețelelor neuronale folosind
alpha_beta_crown și NeuralSat pentru
benchmark-ul cGan al competiției VNN-Comp
2023

Diaconu Laura
Domșa Emanuel
Lapedulce Anastasia
Morariu Ioana-Alexandra
Romanet Rareș

Abstract

In this paper, we attempted to reproduce the results from the VNNCOMP-2023 competition, specifically focusing on the alpha-beta-CROWN and Marabou tools applied to the cGAN benchmark within the same competition. The paper begins with a brief description of how cGAN neural networks operate, followed by a characterization of the dataset and the steps taken to install the tools and run them. We analyzed the data obtained from the runs and compared them with those obtained in the competition. **TO DO: Maybe add a brief idea of the conclusions made after analysis**

Contents

1.	Introducere - Funcționarea rețelei neuronale	2
2.	Caracterizarea setului de date	4
3.	Instalarea și configurarea tool-urilor	5
3.1	alpha-beta-CROWN	5
3.2	NeuralSAT	5
4.	Rularea tool-urilor	5
4.1	alpha-beta-CROWN	5
4.2	NeuralSAT	5
5.	Interpretarea rezultatelor	5
5.1	alpha-beta-CROWN	5
5.2	NeuralSAT	5
5.3	Compararea rezultatelor obținute pe cele două tool-uri . . .	5
6.	Rețelele neuronale în simularea provocărilor reale	5
7.	Concluzii	6
Bibliography		7

1. Introducere - Funcționarea rețelei neuronale

Pentru a putea înțelege mai bine cum funcționează rețeaua neuronală din cadrul **Benchmark-ul cGAN al competiției VNN-Comp 2023** este necesar să înțelegem cum funcționează în general o rețea de tipul GAN, iar apoi una de tip cGAN.

- **GAN**(Generative Adversarial Networks) este un model neuronal mai special [1] a cărui concept poate fi reprezentat ca un joc între două rețele neuronale distincte adversare.
- **cGAN**(Conditional GAN) [2] ghidează procesul de creare a datelor prin încorporarea unor etichete specifice în GAN ca cele două rețele neuronale adversare să se poată orienta după acestea.

O reprezentare grafică a rețelei neuronale din cadrul Benchmark-ul cGAN al competiției VNN-Comp 2023 ar arăta în felul următor:

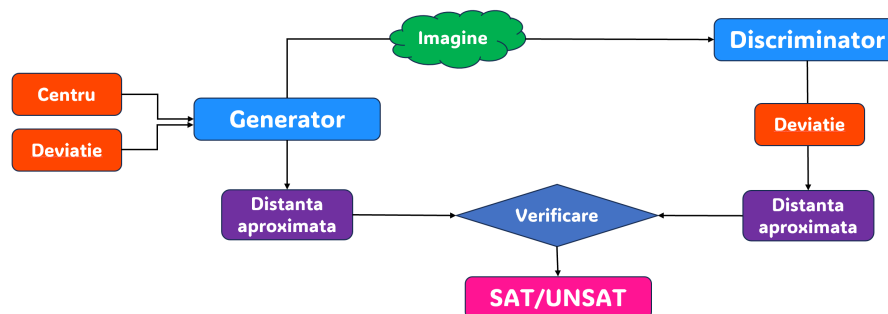


Figure 1: Reprezentare grafică a rețelei neuronale din cadrul Benchmark-ul cGAN al competiției VNN-Comp 2023

Conform graficului, Fig.1, în cazul acestei rețele neuronale cele două rețele distincte adversare se împart în:

- *Generatorul* - pe baza pe datele de intrare(adică etichetele), are rolul de a genera o valoare reprezentată de o distanță și o imagine cu un obstacol aflat la o acea distanță
- *Discriminatorul* - pe baza imaginii returnate de generator, are rolul de a aproxima distanța până la obstacol

Iar datele de intrare sunt:

- *Etichetele* - sunt reprezentate de condiția de distanță și un vector de zgomot, care practic funcționează ca un fel de centru și o deviație

Evaluarea performanței acestei rețele se concentrează pe capacitatea ei de a genera conținut condiționat. În particular, verificarea se bazează pe alinierea distanței prezise de discriminator cu condiția distanței de intrare oferită de generator.

De exemplu, pentru o înțelegere mai clară, am construit următorul exemplu:

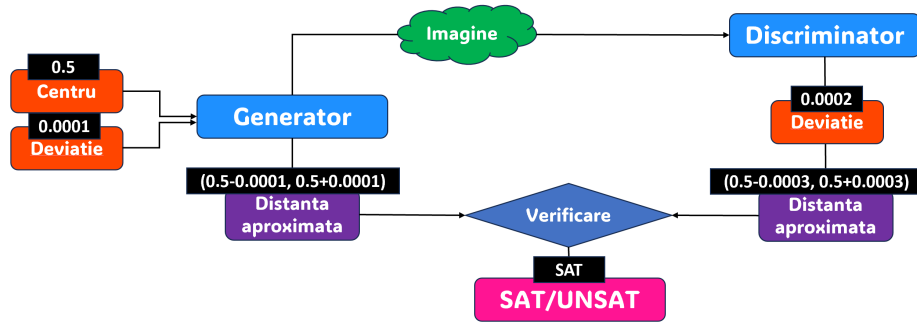


Figure 2: Reprezentare grafică a rețelei neuronale din cadrul Benchmark-ului cGAN al competiției VNN-Comp 2023 - exemplu de date de intrare

În exemplul de mai sus, Fig.2, să presupunem că pentru datele de intrare avem un centru de 0,5 și o deviație de 0,0001. Prin urmare, intervalul distanței la care se poate afla obstacolul este de $(0,5-0,0001, 0,5+0,0001)$.

Obiectivul nostru este să ne asigurăm că distanța prezisă de discriminator se potrivește cu acest interval. Mai exact, distanța prezisă de discriminator trebuie să se situeze în intervalul de intrare adăugând din nou o deviație, de data asta să zicem 0,0002, adică putem obține o distanță aproximativă în intervalul $(0,5-0,0003, 0,5+0,0003)$.

Dacă obținem din partea discriminatorului o distanță din acel interval, acest lucru indică faptul că **imaginile generate respectă condiția de distanță de intrare**.

Rezultatul acestei verificări returnează un rezultat **satisfiabil** dacă se indică o **aproximare corectă a distanței de către discriminator**, în caz contrar se obține un rezultat **nesatisfiabil**.

2. Caracterizarea setului de date

Benchmark-ul cGAN aparține mulțimii de benchmark-uri din cadrul competiției VNN-COMP 2023. Acesta conține o rețea de tip generativă adversarială condiționată și un set de specificații. Benchmark-ul este folosit pentru a verifica corectitudinea și robustețea rețelei pe care o conține. În cadrul set-ului de date, există două tipuri de fișiere: .onnx (Open Neural Network Exchange) și .vnnlib. În fișierele .onnx sunt reprezentate rețele neuronale, acestea conțin informații necesare pentru a executa modelul neuronal. Fișierele .vnnlib conțin specificațiile ce trebuie respectate de rețeaua neuronală, astfel încât aceasta să fie corectă și robustă.

Atât fișierele .onnx cât și .vnnlib au o denumire sugestivă care să indice informații despre conținutul acestora.

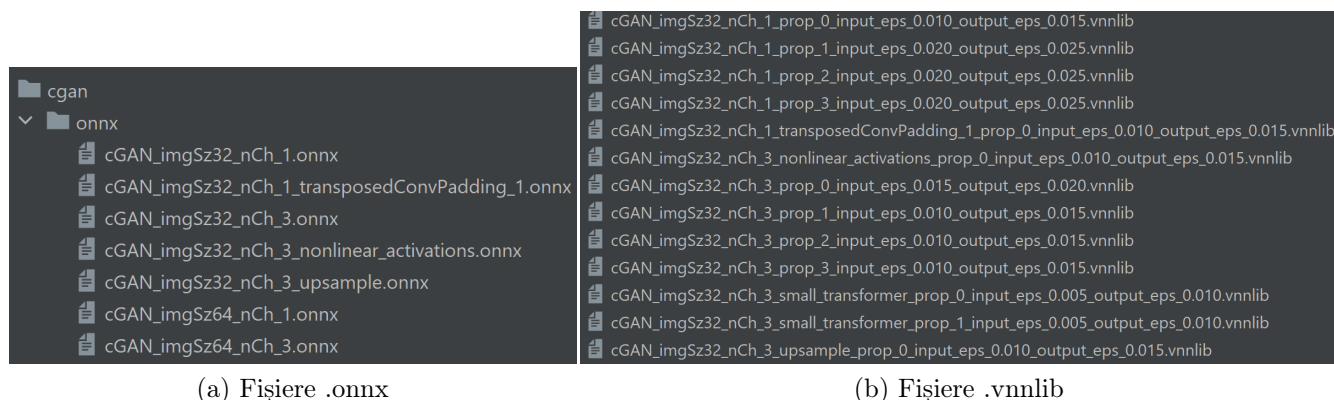


Figure 3: Fișierele benchmark-ului cGAN

- *CGan* - tipul de rețea neurală, în acest caz, o rețea generatoare adversarială condiționată.
- *imgSz32* - dimensiunea imaginii generate, în acest caz, o dimensiune de 32x32 pixeli.
- *nCh_1* - numărul de canale de culoare utilizate în imaginile de intrare sau de ieșire ale rețelei.
- *prop_0* - parametru/proprietate specifică a imaginii
- *input_eps_0* - valoare epsilon utilizată peste datele de intrare ale rețelei.
- *output_eps_0.015* - valoare epsilon utilizată peste datele de ieșire a rețelei.
- *transportedConvPadding_1* - tip specific de convoluție a rețelei.

- *nonlinear_activations* - rețeaua conține funcții de activare non-liniare între straturi.
- *upsample* - modelul neuronal efectuează operații de upsampling, care sunt utilizate pentru a mări dimensiunea spațială a imaginilor sau a datelor. Aceasta poate fi utilă în cazul modelelor generatoare pentru a genera imagini de rezoluție mai mare sau în alte scenarii în care este necesară mărirea dimensiunii datelor.

TO DO: Maybe to be added detailed description of vnnlib file content

...

3. Instalarea și configurarea tool-urilor

3.1 alpha-beta-CROWN

3.2 NeuralSAT

4. Rularea tool-urilor

4.1 alpha-beta-CROWN

4.2 NeuralSAT

5. Interpretarea rezultatelor

5.1 alpha-beta-CROWN

5.2 NeuralSAT

5.3 Compararea rezultatelor obținute pe cele două tool-uri

6. Rețelele neuronale în simularea provocărilor reale

Verificarea rețelelor neuronale este o etapă crucială în dezvoltarea și aplicarea acestor modele avansate. Este esențială din mai multe motive importante.

În primul rând, corectitudinea și fiabilitatea rețelelor neuronale sunt imperative. Ele sunt utilizate într-o varietate de domenii, de la medicină și tehnologie până la securitate cibernetică și vehicule autonome. Verificarea ne asigură că aceste rețele funcționează conform așteptărilor, oferind rezultate precise și fiabile într-o

gamă largă de situații.

Siguranța reprezintă un alt aspect esențial. În domenii critice precum medicina sau industria automotive, erorile în funcționarea rețelelor neuronale pot avea consecințe grave. Verificarea este necesară pentru a identifica și remedia eventualele vulnerabilități care ar putea pune în pericol sistemul.

În industria auto, utilizarea învățării profunde și a viziunii artificiale pentru generarea de imagini noi joacă un rol esențial. Aceste tehnologii permit generarea de imagini sintetice pentru antrenarea și testarea vehiculelor autonome, simularea diverselor scenarii de conducere și optimizarea sistemelor de vizualizare și senzorilor. Ele contribuie la îmbunătățirea siguranței și eficienței în transporturi și la dezvoltarea vehiculelor autonome mai avansate. Utilizarea rețelelor neuronale în generarea de imagini pentru industria auto contribuie prin antrenarea eficientă a algoritmilor, simularea sigură a situațiilor de trafic și optimizarea senzorilor, accelerând dezvoltarea și îmbunătățirea vehiculelor autonome.

De asemenea, verificarea rețelelor neuronale contribuie la prevenirea bias-ului și discriminării. Aceste rețele pot fi influențate de prejudecăți încorporate în datele de antrenament. Prin teste și evaluări riguroase, putem identifica și corecta aceste bias-uri pentru a asigura obiectivitate și corectitudine în rezultatele obținute.

7. Concluzii

Bibliography

- [1] A. Limarc, “What Is a Conditional Generative Adversarial Network? - DZone — dzone.com,” <https://dzone.com/articles/what-is-a-conditional-generative-adversarial-netwo>, 2023.
- [2] M. Mirza and S. Osindero, “Conditional generative adversarial nets,” *CoRR*, vol. abs/1411.1784, 2014. [Online]. Available: <http://arxiv.org/abs/1411.1784>