



**NATIONAL UNIVERSITY OF  
SCIENCE AND TECHNOLOGY  
POLITEHNICA BUCHAREST**



**Doctoral School of Electronics, Telecommunications  
and Information Technology**

**Decision No. 13 from 08-03-2024**

**Ph.D.  
THESIS**

**Andrei GAITA**

---

**METODE DE MACHINE LEARNING PENTRU SUPORTUL  
VERIFICARII CIRCUITELOR INTEGRATE ANALOGICE**

**MACHINE LEARNING METHODS FOR SUPPORTING  
VERIFICATION OF ANALOG INTEGRATED CIRCUITS**

---

**THESIS COMMITTEE**

<b>Prof. Dr. Ing. Mihai CIUC</b> Universitatea Națională de Știință și Tehnologie POLITEHNICA Bucuresti	President
<b>Prof. Dr. Ing. Corneliu BURILEANU</b> Universitatea Națională de Știință și Tehnologie POLITEHNICA Bucuresti	PhD Supervisor
<b>Prof. Dr. rer. nat. Georg PELZ</b> Infineon Technologies AG & Univ. Duisburg-Essen	Referee
<b>Prof. Dr. Ing. Marina TOPA</b> Universitatea Tehnică din Cluj-Napoca	Referee
<b>Prof. Dr. Ing. Cosmin POPA</b> Universitatea Națională de Știință și Tehnologie POLITEHNICA Bucuresti	Referee

**BUCHAREST 2024**

---



# Table of contents

<b>List of tables</b>	<b>v</b>
<b>List of figures</b>	<b>vii</b>
<b>1 Introducere</b>	<b>1</b>
1.1 Asistarea verificării circuitelor integrate analogice și a testelor de producție pe plachete . . . . .	2
1.2 Domeniul de aplicare a cercetării . . . . .	2
1.3 Motivație . . . . .	3
1.4 Structura tezei . . . . .	4
<b>2 Lucrări conexe și fundamente teoretice</b>	<b>5</b>
2.1 Metodologia de verificare a circuitelor integrate analogice . . . . .	5
2.2 Optimizarea verificării prin învățare automată . . . . .	6
2.3 Extragerea caracteristicilor pentru serii de timp . . . . .	6
2.4 Gruparea formelor de undă . . . . .	7
2.4.1 K-means . . . . .	7
2.4.2 Gruparea ierarhica . . . . .	7
2.4.3 Algoritmul DBSCAN . . . . .	8
2.5 Reducerea dimensionalității spațiului caracteristic . . . . .	8
2.6 Evaluarea performanțelor de grupare . . . . .	8
2.6.1 Evaluare externă . . . . .	9
2.6.2 Evaluare internă . . . . .	9
2.7 Verificarea fabricației de IC . . . . .	10
<b>3 Asistarea verificării circuitelor integrate analogice pre-silicon prin intermediul unui algoritm bazat pe SIFT</b>	<b>11</b>
3.1 Introducere . . . . .	11
3.2 Descriptor a caracteristicilor invariante la scală . . . . .	11
3.2.1 Prezentare generală . . . . .	11
3.2.2 Detectarea punctelor cheie critice în seriile de timp IC analogice	12

## Table of contents

3.2.3	Descriptor menit să descrie evenimentele semnificative pentru semnalele IC analogice . . . . .	12
3.2.4	Reprezentarea optimă a sacului de cuvinte . . . . .	14
3.2.5	Gruparea semnalelor . . . . .	14
3.2.6	Vizualizarea spațiului de caracteristici . . . . .	15
3.3	Rezultate experimentale . . . . .	15
3.3.1	Baza de date . . . . .	15
3.3.2	Cazuri de utilizare experimentală a grupării de semnale . . . . .	15
3.4	Concluzii . . . . .	16
<b>4</b>	<b>Tehnici de grupare pentru verificarea analogică IC post-silicon</b>	<b>17</b>
4.1	Introducere . . . . .	17
4.2	Extracția caracteristicilor DTW . . . . .	17
4.2.1	Prezentare generală . . . . .	17
4.2.2	Construcția spațiului caracteristic al semnalelor analogice . . . . .	18
4.3	Model de rețea neuronală pentru semnale analogice . . . . .	18
4.3.1	Prezentare generală . . . . .	18
4.3.2	Model de rețea neuronală convoluțională . . . . .	19
4.3.3	Optimizarea hiperparametrilor autoencoderului . . . . .	19
4.4	Analiza metricilor de grupare . . . . .	20
4.4.1	Overview . . . . .	20
4.4.2	Sensitivity Analysis of Internal Evaluation Metrics . . . . .	20
4.5	Baza de date . . . . .	20
4.6	Cazuri de utilizare experimentală a grupării semnalelor . . . . .	21
4.7	Rezumat și concluzii . . . . .	21
<b>5</b>	<b>Asistarea verificării fabricației de wafere</b>	<b>23</b>
5.1	Prezentare generală . . . . .	23
5.2	Baza de date . . . . .	23
5.3	Detectarea eșecurilor la wafere . . . . .	24
5.4	Rezultate experimentale . . . . .	25
5.5	Concluzii . . . . .	26
<b>6</b>	<b>Concluzii generale</b>	<b>27</b>
6.1	Obiective generale și rezultate . . . . .	27
6.2	Contribuții originale . . . . .	29
6.3	Lista publicațiilor originale . . . . .	29
6.4	Lista rapoartelor tehnice . . . . .	30
6.5	Lucrări viitoare . . . . .	30
	<b>References</b>	<b>31</b>

# List of tables

4.3	Hiperparametrii modelului CNN-AE găsiți după optimizarea bayesiană [GDB <sup>+</sup> 23] . . . . .	19
4.4	Comparison between several clustering metrics . . . . .	20
4.5	Comparație între rezultatele SIFT, CNN-AE și DTW . . . . .	22
5.1	Clasamentul primilor 20 de senzori pe baza preciziei SVM și a metricii Davies-Bouldin . . . . .	25



# List of figures

3.1	Diagrama bloc funcțională de extragere a caracteristicilor invariante la scală [GND <sup>+</sup> 20] . . . . .	13
3.8	Rezultatele grupării folosind metrica de puritate . . . . .	16





# Chapter 1

## Introducere

Verificarea circuitelor integrate (CI) analogice este o componentă majoră în dezvoltarea circuitelor analogice, deoarece asigură că circuitele vor funcționa conform așteptărilor. De asemenea, este o etapă de inspecție esențială, deoarece ajută la identificarea și rectificarea oricăror defecte sau erori introduse în timpul procesului de proiectare.

Verificarea circuitelor integrate analogice constă în două părți. În faza inițială, se examinează conformitatea cu toate cerințele. Funcționalitatea acestor circuite este evaluată pe baza unei varietăți de parametri operaționali, urmată de o analiză comparativă în raport cu repere de performanță prestabilite. Pentru atingerea acestui obiectiv pot fi utilizate diverse metode, cum ar fi simularea în pre-silicon și măsurarea în post-silicon. Aceste metode vor genera o serie de semnale care vor fi salvate pentru o examinare ulterioară.

În cea de-a doua parte a verificării circuitelor integrate analogice, semnalele trebuie verificate vizual de către experți. Acest lucru este necesar deoarece, chiar dacă pentru un anumit produs au fost îndeplinite toate cerințele, există posibilitatea ca acesta să nu funcționeze conform așteptărilor. Provocarea în această etapă a verificării constă în faptul că trebuie să verificăm manual un număr foarte mare de semnale, deoarece trebuie să luăm în considerare semnale din multe combinații posibile de condiții de funcționare.

În lucrarea de față, propunem o metodă pentru a face mai eficient procesul de verificare manuală a semnalelor. Acest lucru este realizat prin procesul de grupare a semnalelor, care permite analiza unui volum vast de date fără a fi nevoie să se examineze individual fiecare semnal. Scopul acestei proceduri de grupare nu este de a verifica cerințele, ci mai degrabă de a inspecta vizual semnalul pentru a confirma absența oscilațiilor neașteptate sau a altor forme de erori.

Obiectivul principal al acestei metode de optimizare este de a reduce timpul necesar pentru verificarea vizuală a semnalelor analogice ale circuitelor integrate, în timp ce un al doilea obiectiv este de a sprijini fabricarea circuitelor integrate prin optimizarea procesului de analiză a semnalelor senzorilor de testare a producției.

## **1.1 Asistarea verificării circuitelor integrate analogice și a testelor de producție pe plachete**

În timpul verificării pre-silicon, se utilizează simulatoare pentru a se asigura că proiectul îndeplinește toate cerințele. De asemenea, este necesar să se evalueze cât mai multe condiții de funcționare pentru a reduce posibilitatea apariției unor defecte de proiectare. În urma acestor simulări, se generează un număr foarte mare de semnale, care trebuie să fie inspectate vizual de către un expert. Cu toate acestea, datorită faptului că scenariile pot fi create automat în cadrul simulării, verificarea devine mai dificilă din cauza creșterii numărului de date care necesită inspecție vizuală. Din cauza numărului mare de semnale care trebuie analizate, gruparea semnalelor poate avea un impact semnificativ în această etapă.

În etapa de verificare post-siliciu, circuitul integrat a fost validat prin simulare și trebuie să fie validat și în condiții de laborator. Măsurarea este un alt proces crucial pentru verificarea circuitelor integrate analogice, deoarece permite proiectanților să măsoare efectiv parametrii circuitelor și să îi compare cu parametrii simulați care au fost anticipați. Semnalele colectate în laborator sunt mai dificil de evaluat decât cele utilizate în mediile de simulare, din cauza prezenței zgomotului și a posibilității apariției unor defecte.

În timpul fabricării waferelor, proiectarea este validată atât înainte cât și după fabricarea acestora; cu toate acestea, este necesar să se testeze circuitele imprimare pe plachete în funcție de parametri specifici pentru a se asigura că producția a funcționat corect. Prin urmare, producția este monitorizată cu ajutorul unei serii de senzori care sunt ulterior analizați pentru a determina dacă procesul de fabricare a funcționat sau nu corect. Mai multe waferi vor fi afectate în cazul în care apare o anumită problemă la producție. Analiza senzorilor și identificarea acelor senzori care sunt corelați cu eroarea de producție reprezintă pasul inițial în identificarea problemei. Datorită faptului că acest proces necesită o analiză laborioasă a semnalelor, ne propunem să oferim sprijin pentru fabricarea de circuite integrate prin implementarea unei metode automate de identificare a senzorilor care sunt corelați cu eroarea de fabricație.

Sprijinirea verificării circuitelor integrate analogice și a testelor de fabricare a waferelor este esențială pentru reducerea muncii manuale implicate. Prin urmare, propunem o nouă metodă de analiză a acestor semnale, oferind o modalitate mai compactă de vizualizare a semnalelor sub formă de clustere și reducând numărul de semnale care trebuie să fie evaluate.

## **1.2 Domeniul de aplicare a cercetării**

Obiectivul acestei lucrări este de a ajuta procesul de verificare a circuitelor integrate analogice și de testare a liniei de producție prin optimizarea muncii manuale cu ajutorul

tehnicilor de învățare automată. În procesul de optimizare a proceselor de verificare, punerea în aplicare a unor astfel de metode poate prezenta anumite provocări:

- Pentru ca algoritmi de învățare automată să fie supuși la antrenare și evaluare, este necesară o cantitate substanțială de date. Sarcina în cauză poate implica achiziția și etichetarea unor cantități mari de date de la diverse circuite integrate și scenarii operaționale în cadrul verificării analogice. O parte din efortul depus în cadrul acestor proiecte a fost dedicată procesului de etichetare a semnalelor. Această procedură a avut ca obiectiv gruparea semnalelor pe baza formelor de undă ale acestora.
- Acuratețea evaluării rezultatelor este direct corelată cu calitatea datelor care sunt utilizate pentru formarea și testarea algoritmilor. Obiectivul actual ar putea necesita o selecție și o prelucrare meticuloasă a datelor pentru a atenua în mod eficient impactul factorilor inutili, cum ar fi zgomotul și alte surse de imprecizie.
- Procesul de selectare a trăsăturilor poate ridica probleme, deoarece poate fi dificil să se stabilească în avans caracteristicile care sunt relevante pentru aplicația specifică. Prin urmare, identificarea combinației optime de caracteristici care trebuie utilizate poate necesita o examinare și o experimentare aprofundată. Pentru a realiza acest lucru, a fost necesar să se evalueze o varietate de tehnici de extragere a caracteristicilor pentru a determina care dintre acestea sunt capabile să extragă informații valoroase din semnalele analogice ale circuitelor integrate pentru a fi utilizate ulterior în procesul de grupare.
- Sarcina de a selecta și adapta un algoritm de învățare automată adecvat pentru această aplicație specială poate reprezenta o provocare semnificativă. Acest lucru se datorează faptului că există mai multe tehnici și proceduri care ar putea fi optimizate, fiecare având caracteristici specifice. Mai multe procese care au potențial de optimizare includ identificarea evenimentelor de semnal neașteptate, identificarea senzorilor care reflectă defectarea cipului și susținerea fabricării de circuite integrate.

Utilizarea tehnicilor de învățare automată pentru a reduce cantitatea de inspecție manuală necesară în procesul de verificare analogică poate fi dificilă, dar are, de asemenea, potențialul de a oferi beneficii considerabile în ceea ce privește acoperirea și eficiența verificării circuitelor integrate analogice.

### 1.3 Motivație

Utilizarea metodelor de învățare automată în scopul sprijinirii proceselor de verificare analogică are potențialul de a reduce volumul de muncă manuală, facilitând în același

timp analiza unor cantități mari de date legate de scenarii operaționale complexe. Acest lucru poate contribui la accelerarea întregului proces de proiectare a circuitelor integrate, îmbunătățind în același timp eficiența procesului de verificare. În consecință, există multiple justificări pentru utilizarea tehnicilor de învățare automată pentru a optimiza procesul de verificare:

Obiectivul acestui studiu este de a maximiza optimizarea întregii proceduri de dezvoltare a circuitelor integrate analogice prin integrarea și adaptarea metodologiilor de învățare automată în procesele de verificare a circuitelor integrate, cu scopul de a obține cel mai mare impact posibil. Utilizarea tehnicilor de învățare automată pentru îmbunătățirea și automatizarea procesului de verificare a circuitelor integrate analogice are potențialul de a aduce beneficii semnificative.

## 1.4 Structura tezei

În capitolul 2, prezentăm elementele fundamentale ale învățării automate, inclusiv tehnicile de grupare utilizate în procesul de optimizare a verificării, tehnicile de învățare automată și stadiul actual al tehnicilor de optimizare bazate pe învățarea automată pentru verificarea circuitelor integrate analogice.

În capitolul 3, vom prezenta o metodă de îmbunătățire a procesului de verificare pre-silicon prin proiectarea unui algoritm de grupare a semnalelor pe baza unor forme de unda bine cunoscute identificate de inginerii de verificare. Această tehnică de grupare și optimizare constă într-un algoritm de extragere a caracteristicilor invariante în timp, un algoritm de grupare și un algoritm de vizualizare a semnalelor în spațiul caracteristicilor.

Capitolul 4 pune accentul pe evenimentele prezente în semnalele post-silicon pe care algoritmul prezentat în capitolul 3 nu le poate caracteriza. În acest cadru, am dezvoltat două metode de extragere automată a caracteristicilor: una bazată pe rețele neuronale și cealaltă pe o abordare de procesare a semnalelor numită Dynamic Time Warping. Tot în acest capitol, am efectuat o examinare comparativă a mai multor metrici de grupare în legătură cu cazul nostru de utilizare a verificării circuitelor integrate analogice.

În capitolul 5 ne concentrăm pe sprijinirea fabricării circuitelor integrate pe wafere. Acest lucru este realizat utilizând cea mai eficientă tehnică de extragere a caracteristicilor prezentată în capitolul 4 în combinație cu mai mulți algoritmi de clasificare și metrici de grupare optime pentru cazul de utilizare specific al semnalelor de circuite integrate analogice.

În capitolul 6, formăm o concluzie generală cu privire la abordările care au fost propuse, discutăm influența asupra problemei actuale și evidențiem contribuțiile importante pe care autorul le-a adus. În concluzia tezei, sunt prezentate câteva sugestii pentru potențiale cercetări viitoare pe tema îmbunătățirii eficienței verificării circuitelor integrate analogice.

# Chapter 2

## Lucrări conexe și fundamente teoretice

Acest studiu se va concentra pe asistarea verificării circuitelor integrate analogice pentru a reduce efortul manual implicat în acest proces. Odată cu evoluția industriei semiconductoarelor, tehnologia permite mai multe cerințe, iar procedura de verificare a devenit din ce în ce mai laborioasă [GS19]. Acest lucru s-a datorat în principal componentei umane a sistemului de verificare, care nu poate ține pasul cu cantitatea enormă de date care trebuie evaluate vizual [CK07].

Verificarea manuală se referă la procedura de evaluare a eficacității operaționale a circuitelor analogice prin intermediul competenței umane sub formă de examinare vizuală. În etapa de pre-silicon a procesului de proiectare, care are loc înainte ca circuitul integrat să fie efectiv construit, verificarea manuală este frecvent utilizată pentru a detecta comportamente neașteptate în diverse situații de funcționare [GXGM19]. Circuitele analogice pot fi simulate cu ajutorul aplicațiilor software, iar rezultatele pot fi comparate cu parametrii de performanță prevăzuți.

Îmbunătățirea și facilitarea verificării manuale poate influența semnificativ procesul de verificare analogică a circuitelor integrate, atât în faza pre-silicon, cât și în cea post-silicon. Motivul este că examinarea manuală efectuată de oameni poate necesita o investiție semnificativă atât de timp, cât și de resurse. Prin urmare, prin optimizarea acestui aspect particular al procesului de verificare, este posibil să se îmbunătățească eficiența procesului general de proiectare a circuitelor integrate analogice.

### 2.1 Metodologia de verificare a circuitelor integrate analogice

Obiectivul principal al metodologiei de verificare a circuitelor integrate analogice se referă la examinarea semnalelor simulate sau măsurate ale circuitelor integrate analogice. Semnalele avute în vedere sunt în principal serii de timp unidimensionale care respectă criteriile specifice, inclusiv timpul de stabilizare, nivelurile de depășire sau de scădere

și valorile semnalului. În plus, în aceste semnale se caută în mod deliberat fenomene neobișnuite, cum ar fi modelele oscilatorii sau morfologia atipică a semnalelor.

Folosirea unor instrumente automatizate care sunt capabile să evalueze datele simulate și să identifice orice abateri de la standardele de performanță stabilite poate fi benefică pentru eficientizarea procesului de inspecție manuală [MGK<sup>+</sup>05]. Îmbunătățirea calității vizualizării datelor servește ca măsură suplimentară la automatizare. În situațiile în care un volum mare de date necesită o evaluare manuală, inginerii pot utiliza o tehnică de vizualizare a datelor care să permită vizualizarea și compararea simultană a sute de semnale într-un mod clar și ușor de înțeles.

## 2.2 Optimizarea verificării prin învățare automată

Procedura de verificare a circuitelor integrate analogice poate fi eficientizată prin utilizarea diferitelor strategii de învățare automată. Aceste metode implică algoritmi de formare și de adaptare pentru a extrage caracteristicile relevante din semnalele analogice. Primul pas al utilizării tehnicilor de învățare automată pentru verificarea circuitelor integrate analogice este colectarea și adnotarea unor volume mari de date de la o varietate de produse și situații de funcționare. Aceste date pot fi apoi folosite pentru a dezvolta tehnici de învățare automată pentru a descoperi modele specifice în semnalele analogice [MCC<sup>+</sup>22].

Cercetătorii au folosit abordări de învățare automată nesupravegheate în [YWCW21]. Scopul studiului a fost acela de a evalua semnalele și de a găsi comportamente defectuoase direct din seriile de timp, fără niciun pas de preprocesare. Constatările au demonstrat că abordările de învățare automată au fost capabile să identifice în mod eficient comportamente defectuoase în cadrul semnalelor, ceea ce a dus la o creștere a vitezei și a preciziei procesului de inspecție vizuală.

## 2.3 Extragerea caracteristicilor pentru serii de timp

Procesul de obținere a caracteristicilor sau trăsăturilor relevante dintr-o bază de date pentru a le include într-un model de rețea neuronală sau pentru a le utiliza într-un algoritm este denumit extracție de caracteristici și este unul dintre primii pași posibili implicați în învățarea automată [KD16]. Scopul extragerii caracteristicilor este de a localiza acele aspecte ale datelor care sunt cele mai semnificative și pertinente pentru problema în cauză, ajutând astfel modelul să clasifice corect datele.

Când se referă la prelucrarea semnalelor, termenul "extragerea caracteristicilor" se poate referi la procesul de extragere a aspectelor unui semnal, cum ar fi spectrul de frecvență, forma formele de undă sau calitățile statistice ale semnalului [RHW<sup>+</sup>16]. După aceea, aceste caracteristici pot fi trimise într-un model de învățare automată sub

formă de intrare, unde modelul poate fi antrenat pentru a clasifica sau prezice semnalul pe baza acestor caracteristici.

## 2.4 Gruparea formelor de undă

Gruparea formelor de unda este o tehnică de învățare automată care presupune formarea de clustere din seturi de puncte de date care prezintă similitudini. Obiectivul principal al grupării este de a identifica modelele recurente din cadrul datelor. Acest proces le permite analiștilor să obțină informații noi și să obțină o înțelegere mai cuprinzătoare a distribuției datelor, facilitând chiar și vizualizarea [XW10].

Gruparea formelor de unda este o metodă valoroasă pentru a obține cunoștințe despre distribuția de bază a datelor, precum și pentru a localiza modelele și corelațiile prezente în cadrul acestora. Datorită acestor factori, considerăm că gruparea poate avea un impact semnificativ asupra optimizării procesului de verificare a circuitelor integrate analogice, având în vedere cantitatea mare de date implicate.

### 2.4.1 K-means

K-means este un algoritm care împarte un set de date într-un număr predeterminat de grupuri în funcție de distanța dintre punctele de date. Obiectivul lui k-means este de a minimiza suma pătratelor în cadrul unui grup, o măsură a distanței dintre punctele de date din interiorul fiecărui cluster [JJJ<sup>+</sup>20].

Faptul că gruparea k-means poate fi implementată rapid și cu puțin efort este unul dintre numeroasele puncte forte ale metodei. Algoritmul de grupare K-means este o metodă puternică și populară care împarte un set de date în grupuri pe baza distanței medii dintre fiecare pereche de puncte de date din setul de date. Găsirea de modele și corelații în date, precum și obținerea de informații despre structura de bază a setului de date sunt, de asemenea, aplicații eficiente ale acestei tehnici.

### 2.4.2 Gruparea ierarhica

Gruparea ierarhică este o tehnică care grupează punctele de date pe baza unei reprezentări a spațiului caracteristic și creează o organizare stratificată a datelor sub formă de arbore. Acest tip de grupare poate fi împărțit în două categorii principale: aglomerativă (sau ascendentă) și divizivă (descendentă) [PSJ15]. Cea mai populară metodă de grupare se numește grupare ierarhică aglomerativă și implică începerea cu un număr mare de grupe mici și apoi fuzionarea lor pe baza asemănărilor dintre ele până când toate punctele de date sunt incluse într-un singur grup. În schimb, gruparea ierarhică divizivă începe cu un singur cluster mare și apoi îl împarte în mai multe clustere mai mici, în funcție de similaritatea dintre punctele de date din fiecare dintre clusterelor mai mici, până când fiecare punct de date se află în propriul cluster.

Gruparea ierarhică generează o reprezentare vizuală a grupurilor sub forma unei dendrograme, care oferă o interpretare rezonabil de cuprinzătoare, în special pentru datele analogice de verificare IC.

### 2.4.3 Algoritmul DBSCAN

Abordarea cunoscută sub numele de Density-Based Spatial Clustering of Applications with Noise (DBSCAN) propusă de autori în [Den20] este utilizată pentru a grupa punctele de date individuale incluse într-un set de date. Ea se bazează pe ipoteza că grupurile dintr-un set de date sunt produse de regiuni de puncte cu densitate mare, care sunt apoi separate de regiuni cu densitate mică. DBSCAN necesită definirea a 2 hiperparametri; primul este Epsilon, distanța maximă față de cei mai apropiați vecini, iar al doilea este MinPts, numărul minim de puncte necesare pentru ilustrarea unui cluster [BY21].

Capacitatea algoritmului de a număra automat numărul de grupuri prezente în setul de date, precum și capacitatea sa de a evidenția punctele individuale care nu aparțin niciunui grup, sunt două dintre numeroasele puncte forte ale algoritmului. În plus, acesta este relativ eficient, deoarece, pentru a-și face treaba, trebuie să calculeze doar distanța dintre punctele care se află la o distanță de Eps unul față de celălalt.

## 2.5 Reducerea dimensionalității spațiului caracteristic

Vizualizarea semnalelor într-o manieră concisă și inteligibilă prin utilizarea caracteristicilor extrase algoritmic este un element crucial al acestui studiu. Având în vedere că spațiul caracteristic este caracterizat de mai multe dimensiuni, este necesar să se utilizeze tehnici de reducere a dimensionalității pentru a facilita vizualizarea semnalelor ca puncte într-un spațiu bidimensional.

Reducerea dimensionalității, denumită și reducerea dimensionalității spațiului caracteristic, este o tehnică utilizată în domeniul învățării automate pentru a reduce numărul de caracteristici sau dimensiuni într-un set de date dat, păstrând în același timp cantitatea maximă de informații pertinente. Acest lucru aduce beneficii din numeroase motive, inclusiv din următoarele:

Procesul de reducere a dimensionalității necesită o analiză atentă a compromisurilor sale, deoarece poate reduce potențial cantitatea de informații disponibile pentru model, ceea ce poate compromite performanța acestuia.

## 2.6 Evaluarea performanțelor de grupare

Monitorizarea eficienței performanțelor de grupare este foarte importantă. Putem împărți algoritmi de evaluare a grupării disponibili în două mari categorii: metrici care necesită cunoașterea adevărului de bază, numite metrici de evaluare externă [dSCF<sup>+</sup>12], și metrici



care evaluează rezultatul grupării în sine, numite metrice de evaluare internă [LLX<sup>+</sup>13]. Fiecare dintre aceste tipuri are avantaje și dezavantaje; prin urmare, avem următoarele calități pentru fiecare categorie.

### 2.6.1 Evaluare externă

#### Metrica de puritate

Puritatea este una dintre măsurile utilizate pentru validarea rezultatelor, fiind o statistică ce indică gradul în care un grup cuprinde doar un singur tip de semnal [RJ18]. Această măsură este substanțial mai eficientă atunci când este vorba de un număr mic de grupuri de semnale, deoarece este mult mai ușor să se obțină cel mai înalt nivel de puritate a semnalului atunci când există numeroase grupuri mici de semnale. Dacă alegem să asociem fiecare punct cu un cluster, vom obține un scor de puritate de sută la sută, ceea ce face ca această statistică să fie lipsită de sens. În selectarea acestei metrice, am ținut astfel cont de necesitatea de a restricționa numărul de cluster și de a avea date echilibrate, deoarece metrica de puritate nu produce rezultate relevante în cazul unor date dezechilibrate.

#### Indexul Fowlkes-Mallows

Metrica de evaluare externă Fowlkes-Mallows (FM) este utilizată pentru a măsura capacitatea procesului de grupare de a reproduce distribuția manuală a adnotărilor [ISAS21]. Metrica FM are o valoare care indică modul în care rezultatele grupării se potrivesc cu adnotările manuale. Pentru a determina indicele Fowlkes-Mallows, propus în [FM83].

Această statistică este bună datorită faptului că se obțin rezultate semnificative atât pentru seturile de date dezechilibrate, cât și pentru cele care conțin valori aberante, deoarece sunt incluse atât precizia, cât și reamintirea. Având în vedere exhaustivitatea sa ca metrică de evaluare externă în comparație cu puritatea, această metrică este, de asemenea, potrivită pentru compararea rezultatelor grupării cu adnotările manuale efectuate pe semnale achiziționate de la circuite integrate analogice.

### 2.6.2 Evaluare internă

Scopul categoriei de măsurători de evaluare internă este de a determina gradul în care semnalele pot fi diferențiate unele de altele în spațiul caracteristic. Este esențial să luăm în considerare acest aspect, deoarece este mijlocul prin care putem afirma în mod clar că un algoritm este capabil să extragă caracteristici mai relevante decât altul [LLX<sup>+</sup>10]. Deși aceste metrice evaluează în esență aceleași aspecte, cum ar fi gradul de compactitate a clusterelor și distanța dintre cluster [VK16], fiecare dintre ele are caracteristici unice care o fac mai mult sau mai puțin ideală pentru gruparea semnalelor analogice. Având în

vedere acest lucru, este necesară o comparație obiectivă a acestor măsuri, ținând cont de ceea ce dorim să punem în evidență în optimizarea verificării semnalelor și a detectării defectelor.

### **Metrica Davies-Bouldin**

Metrica Davies-Bouldin este un indicator comun de performanță a grupării utilizat în literatura științifică pentru a evalua eficacitatea abordărilor de grupare [SMMS20]. Acest algoritm examinează întreaga calitate a grupării, inclusiv gradul de compactizare a punctelor, precum și distanța dintre grupurile de puncte. Prin urmare, clusterelor care sunt mai îndepărtate unele de altele și mai dense vor avea un scor mai mic, indicând o performanță superioară [KMH17].

### **Coeficientul Silhouette**

Sensibilitatea coeficientului de silhouette la numărul de clusterelor necesită compararea coeficientului de silhouette la diferite numere de clusterelor pentru a stabili numărul optim de clusterelor. Este demn de remarcat faptul că coeficientul de silhouette este sensibil la scara datelor, ceea ce necesită normalizarea datelor înainte de calcularea coeficientului de silhouette [SN20].

### **Indexul Calinski-Harabasz**

Indicele Calinski-Harabasz (indicele CH) prezentat în [CH74], denumit, de asemenea, criteriul raportului de varianță, este o măsură care cuantifică gradul de compactitate și de separare a clusterelor într-un anumit set de date. Utilizarea unui criteriu de evaluare este viabilă în scopul stabilirii celui mai potrivit număr de clusterelor într-un algoritm de grupare.

## **2.7 Verificarea fabricației de IC**

În timpul fabricării waferelor, numeroși senzori monitorizează procesul de producție. În cazul unei probleme de producție, acești senzori trebuie să fie analizați atât automat, cât și manual pentru a determina cauza potențială. Procesul de analiză a senzorilor de plachete pentru a identifica problemele poate fi îmbunătățit în diverse moduri prin utilizarea învățării automate [CTGK22]. Una dintre aplicațiile învățării automate este automatizarea colectării și analizei datelor, ceea ce poate spori eficiența acestor proceduri. Algoritmii de învățare automată pot fi utilizați pentru a identifica cauzele fundamentale ale unei probleme prin analiza unor cantități mari de date legate de proiectarea circuitului integrat, procesul de fabricație și rezultatele testelor [RZWD15]. Acest lucru poate facilita accelerarea procesului de investigare a cauzei fundamentale a problemei, diminuând astfel cantitatea de timp și resurse necesare pentru a determina sursa problemei.

# Chapter 3

## Asistarea verificării circuitelor integrate analogice pre-silicon prin intermediul unui algoritm bazat pe SIFT

### 3.1 Introducere

În procesul de verificare pre-silicon a circuitelor integrate analogice, este necesară verificarea vizuală a semnalelor generate în urma simulărilor. Etapa de verificare manuală a semnalelor este necesară deoarece există posibilitatea apariției unor comportamente neașteptate pentru care nu se pot scrie în prealabil verificatori. Acest proces este extrem de laborios, după cum se explică în articolul [KC06]; prin urmare, propunem ca soluție la această problemă tehnici multiple de grupare a semnalelor pentru a facilita evaluarea semnalelor.

Se efectuează mai multe experimente pe fiecare tip de circuit integrat, după care se generează și se analizează semnalele. Această strategie este foarte benefică deoarece, pentru fiecare test individual, semnalele vor fi grupate, permițând inginerilor să se concentreze asupra câtorva grupuri de semnale, mai degrabă decât să inspecteze fiecare semnal în parte. Această tehnică automatizată îmbunătățită de grupare a cantităților mari de date face ca procesul de verificare să fie semnificativ mai rapid și mai fiabil, deoarece, dacă oricare dintre semnalele măsurate sau simulate prezintă evenimente neașteptate, acestea vor fi clar separate de restul semnalelor.

### 3.2 Descriptor a caracteristicilor invariante la scală

#### 3.2.1 Prezentare generală

Una dintre principalele provocări ale grupării formelor de undă este lungimea variabilă a semnalelor. Această problemă apare adesea în cazul semnalelor provenite din simulări.

O altă provocare este reprezentată de semnalele de înaltă frecvență generate de simulare în timpul fazei de pre-siliciu, ceea ce poate duce la provocări în identificarea metricii de similaritate adecvate necesare pentru algoritmul de grupare. Prin urmare, invarianța temporală și a lungimii semnalului este o cerință de bază pentru gruparea formelor de undă a semnalelor de verificare. Realizarea acestui lucru se poate face fie prin alegerea metodelor adecvate de extragere a caracteristicilor.

Procesul de extragere a caracteristicilor implică examinarea semnalului analogic al IC în trei etape succesive pentru a obține o reprezentare multidimensională precisă a caracteristicilor, așa cum se arată în figura 3.1, care descrie diagrama bloc funcțională a algoritmului complet de extragere a caracteristicilor.

Obiectivul algoritmului bazat pe SIFT este de a extrage caracteristicile relevante din semnale, pe baza similitudinii formei de undă. Pentru a extrage aceste caracteristici, algoritmul identifică punctele cheie din semnale pentru a localiza caracteristicile semnificative pentru extragere. Un descriptor este apoi utilizat pentru a caracteriza aceste regiuni cât mai precis posibil, utilizând cât mai puțini coeficienți. Descriptorii sunt obținuți prin efectuarea unei transformări continue de wavelet-uri (CWT), urmată de o transformare discretă în cosinus 2D (DCT) pe scalograma rezultată din CWT. Unul dintre avantajele utilizării CWT este capacitatea sa de a descrie în mod eficient spectrul timp-frecvență din vecinătatea punctelor cheie într-o manieră rezistentă, după cum se menționează în [LO12]. Motivul pentru care am aplicat CWT la semnalele analogice a fost acela de a caracteriza cu precizie comportamentele tranzitorii scurte, deoarece unele erori de semnal pot avea o frecvență ridicată și o durată scurtă. În plus, s-a utilizat DCT pentru a comprima caracteristicile datorită numărului mare de coeficienți din scalograma rezultată [ANR74].

### **3.2.2 Detectarea punctelor cheie critice în seriile de timp IC analogice**

Pentru a începe procesul de extragere a caracteristicilor, pasul inițial implică identificarea punctelor cheie. Aceste puncte cheie sunt definite ca puncte de interes pe baza abordării SIFT. Pentru a identifica punctele de interes menționate mai sus, se va utiliza un filtru gaussian. Acest tip de filtru liniar este utilizat frecvent în procesarea imaginilor pentru a atenua zgomotul sau a estompa imaginile, după cum se explică în lucrarea lui Lowe [Low04] pentru abordarea bazată pe SIFT.

### **3.2.3 Descriptor menit să descrie evenimentele semnificative pentru semnalele IC analogice**

Descriptorul a fost dezvoltat astfel încât să ofere o reprezentare foarte unică a regiunilor din jurul punctelor cheie. Această etapă implică luarea în considerare a două operații

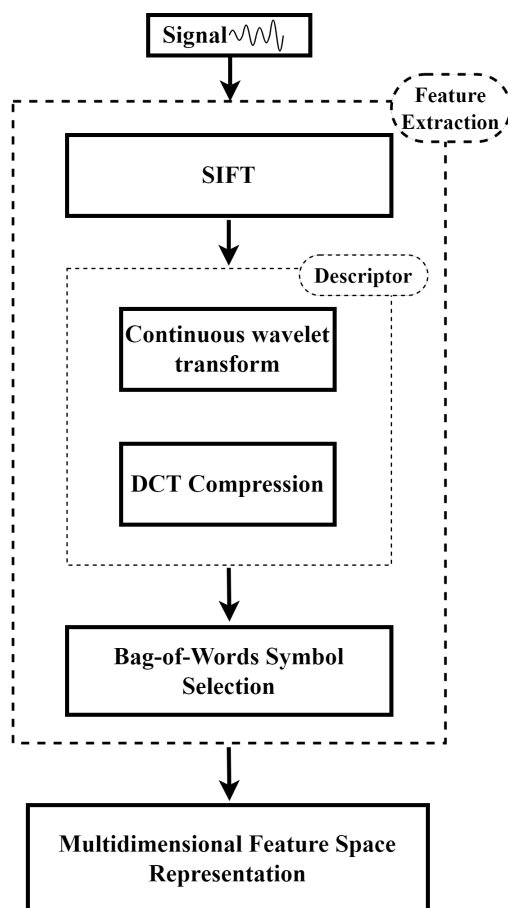


Fig. 3.1 Diagrama bloc funcțională de extragere a caracteristicilor invariante la scală [GND<sup>+</sup>20]

diferite: mai întâi, un bloc de transformare wavelet și apoi o transformare discretă în cosinus (DCT) [ANR74].

Obiectivul CWT este de a reprezenta cu exactitate în domeniul timp-frecvență segmentele de semnal din vecinătatea fiecărui punct cheie. Această reprezentare, cunoscută sub numele de scalogramă, este potrivită pentru a descrie variațiile neperiodice și tranzitorii de scurtă durată, care sunt caracteristici comune ale setului nostru de date. Se va defini un segment de timp pentru analiza de frecvență care urmează să fie efectuată și, deoarece acest lucru se va întâmpla pentru fiecare punct cheie localizat de blocul SIFT, se vor efectua mai multe reprezentări de frecvență pentru fiecare semnal individual, în funcție de numărul de puncte cheie.

Pentru a reprezenta cu acuratețe aparițiile scurte cu frecvență fluctuantă, am utilizat undele Morse generalizate datorită eficienței lor computaționale în examinarea discontinuităților izolate, așa cum se menționează în literatura de specialitate [LO12].

Cu ajutorul acestei noi reprezentări ordonate, este posibilă comprimarea scalogramei prin utilizarea unui număr de coeficienți mai mic de 64. În contextul prezentei lucrări, cercetarea empirică care utilizează datele furnizate a arătat că se observă o eroare de reconstrucție mai mică de 5% în intervalul de coeficienți [25-35]. Coeficienții DCT au

fost cuantificați folosind 31 de niveluri de cuantificare, așa cum este descris în [RS01], pentru a obține un nivel de compresie mai bun și pentru a permite etapa de agregare a caracteristicilor.

Prin urmare, pentru fiecare segment de semnale definit de un punct cheie, 31 de coeficienți vor fi utilizați pentru a caracteriza regiunea respectivă. Acești coeficienți descriu versiunea propusă a descriptorului specific semnalelor analogice. Prin utilizarea analizei de frecvență și a tehnicilor de compresie bazate pe DCT, suntem capabili să descriem evenimente tranzitorii de scurtă durată, care sunt de cea mai mare importanță pentru procesul de verificare.

### 3.2.4 Reprezentarea optimă a sacului de cuvinte

Această fază finală de construire a spațiului caracteristic utilizează generarea unui dicționar de simboluri care este utilizat pentru a descrie semnalele sub formă de histograme de simboluri pe parcursul procedurii de grupare. Această caracterizare a semnalelor (cunoscută și sub denumirea de sac de cuvinte) are avantajul de a asigura invariabilitatea lungimii semnalului și a numărului de puncte cheie. Motivul pentru care am ales această strategie a fost acela de a defini caracteristicile unui semnal pe baza formei semnalului în cadrul unei ferestre centrate pe punctul cheie. Pentru a realiza acest lucru, a fost necesar să identificăm formele de undă cel mai frecvent întâlnite în jurul punctelor cheie și apoi să creăm un dicționar bazat pe aceste forme de undă.

Pentru a crea reprezentarea de tip sac de cuvinte a semnalului din setul de date, este necesar să împărțim spațiul descriptorilor DCT rezultați pe baza separabilității lor în hiperspațiu. Acest lucru se realizează utilizând DBSCAN [KuRA<sup>+</sup>14] pe descriptorii derivați din întregul set de date. În scenariul nostru, cele 31 de clustere rezultate vor constitui dicționarul de simboluri.

### 3.2.5 Gruparea semnalelor

Această etapă este realizată pentru a separa semnalele fiecărui test individual din setul de date în grupuri separate cu propria lor reprezentare a sacului de cuvinte. Acest lucru va permite inginerului de verificare să verifice vizual seturile de semnale mai ușor. În acest scop, am utilizat gruparea k-means [KYY20], care este o abordare simplă și eficientă atunci când există o graniță clară între grupuri în spațiul multidimensional al caracteristicilor. De asemenea, am utilizat gruparea ierarhică cu o abordare aglomerativă [KYY20] din motive de comparație.

K-means a fost ales pentru că grupează cu succes punctele cu caracteristici liniar diferite într-un spațiu multidimensional. În acest caz, abordarea de grupare a fost aplicată spațiului multidimensional al caracteristicilor. Pentru a aplica abordarea k-means, trebuie să furnizăm numărul de clustere. Această valoare a fost aleasă pe baza cantității de etichete din setul de date pentru fiecare test.

### **3.2.6 Vizualizarea spațiului de caracteristici**

Vizualizarea este benefică și atunci când există valori aberante foarte distincte de celelalte semnale, care vor fi afișate în spațiul bidimensional, departe de celelalte locații. Această vizualizare a fost realizată prin intermediul algoritmului de analiză în componente principale (PCA) [MR93], care a avut rolul de a reduce la două dimensiuni spațiul caracteristicilor generate de algoritmul propus.

Datorită faptului că această reducere a dimensionalității are o pierdere de informație [GK12], este vital să se știe câtă informație se păstrează în primele două componente pentru o separare eficientă și o vizualizare 2D. Pentru a stabili adecvarea acestui tip de prezentare pentru semnale analogice într-un spațiu bidimensional, păstrând în același timp separabilitatea multidimensională, a fost imperativ ca setul nostru de date să prezinte o proporție semnificativă de variație în cele două componente principale inițiale [DYL08].

## **3.3 Rezultate experimentale**

### **3.3.1 Baza de date**

În această etapă a cercetării noastre, am creat un set de date care utilizează date provenite din 10 simulări, ceea ce reprezintă un total de 2.950 de semnale. Aceste semnale au fost obținute într-un mediu de simulare folosind regulatoare cu pierderi scăzute de curent. Pentru fiecare test, un inginer de verificare etichetează manual semnalele în două sau trei grupe care reflectă comportamente pozitive distincte și evenimente de eșec. Deoarece efectuăm o clasificare nesupravegheată (grupare), acest lucru ajută, de asemenea, la evaluarea performanței algoritmului.

Lungimea semnalelor diferă considerabil între întregul set de date. Acest lucru justifică necesitatea unui algoritm de grupare care să asigure invarianța lungimii. În plus, setul de date conține semnale care prezintă fenomene relevante pentru verificarea circuitelor integrate analogice.

### **3.3.2 Cazuri de utilizare experimentală a grupării de semnale**

În această lucrare, au fost efectuate o serie de experimente cu scopul principal de a determina măsura în care algoritmul bazat pe SIFT poate extrage caracteristici valoroase din semnalele analogice de IC. În plus, am determinat care este cel mai bun algoritm de grupare aplicat pe spațiul caracteristic produs de algoritmul bazat pe SIFT. Pentru a examina toate aceste aspecte, au fost utilizate zece grupuri de semnale prezentate în capitoul precedent, împreună cu algoritmul SIFT și doi algoritmi de grupare, K-means și grupare ierarhica.

O primă analiză a fost efectuată pe baza metricii purității grupării pentru a vizualiza gradul de separare între semnale pe baza adnotării manuale. În figura 3.8 este prezentată comparația bazată pe puritatea celor doi algoritmi de grupare aplicați aceluiași spațiu de caracteristici. În cazul grupării bazate pe k-means, așa cum se arată în Figura 3.8, am obținut o puritate medie de 97% pentru teste, ceea ce demonstrează o excelentă separabilitate a semnalelor în spațiul de caracteristici. De asemenea, această analiză demonstrează că gruparea ierarhică produce rezultate de grupare mai puțin precise, ceea ce sugerează că organizarea semnalelor sub forma unei dendrograme poate să nu fie soluția optimă pentru scopul acestei lucrări.

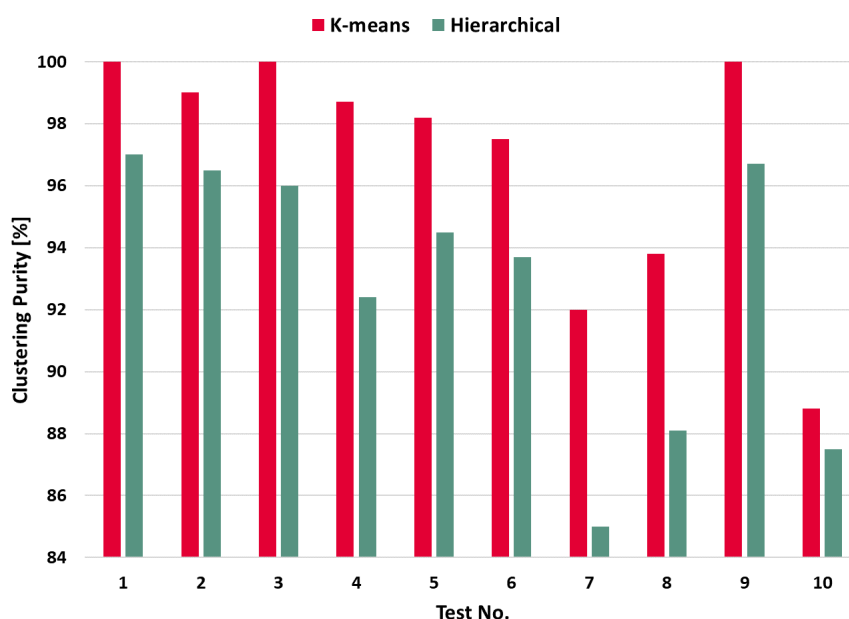


Fig. 3.8 Rezultatele grupării folosind metrica de puritate

### 3.4 Concluzii

În acest capitol, am prezentat o metodă care poate fi utilizată pentru gruparea formelor de undă de testare a circuitelor în funcție de similitudinea lor, în scopul de a reduce timpul necesar pentru verificarea pre-silicon. Am realizat acest lucru prin utilizarea unei abordări de tip sac de cuvinte care utilizează un cadru SIFT pentru extragerea caracteristicilor semnalului. Pe lângă rezultatele foarte bune de grupare, cu o puritate de 98%, această abordare are și avantajul de a obține invarianțe ale scării temporale și ale lungimii semnalului. În ciuda disponibilității seturilor de date adnotate, am dezvoltat un algoritm de învățare nesupravegheată cu scopul de a-l utiliza pe un set de date semnificativ mai mare, în care procesul de etichetare este impracticabil din cauza efortului asociat.



# Chapter 4

## Tehnici de grupare pentru verificarea analogică IC post-silicon

### 4.1 Introducere

În acest capitol, am dezvoltat două metode pentru a ajuta fluxul de verificare: una bazată pe DTW și cealaltă pe autoencoder. Aceste două tehnici pot fi utilizate pentru extragerea caracteristicilor din semnalele IC analogice cu comportamente neașteptate. DTW și autoencoderul sunt două abordări care au potențialul de a fi eficiente atunci când vine vorba de extragerea caracteristicilor din semnalele analogice de circuite integrate care prezintă comportamente neașteptate. Atât caracteristicile datelor, cât și obiectivele analizei vor juca un rol în selectarea algoritmului adecvat. Pentru a determina ce metodă ar da cele mai bune rezultate pentru o aplicație specifică, ar putea fi esențial să se încerce o varietate de algoritmi și combinații.

### 4.2 Extracția caracteristicilor DTW

#### 4.2.1 Prezentare generală

Algoritmul bazat pe DTW este utilizat în tehnica propusă pentru a măsura diferențele de formă de undă între seriile de timp provenite din același test. Pentru a evalua metoda propusă, am utilizat mai multe tehnici metrice de clusterizare pentru a măsura separabilitatea spațiului caracteristic între clustere și timpul de execuție al algoritmului. Această abordare, în forma sa cea mai elementară, implică utilizarea unei tehnici de deformare între 2 semnale pentru a le compara cât mai clar posibil. În urma executării algoritmului, va rezulta coeficientul DTW care indică gradul în care două semnale sunt comparabile între ele. Deoarece tehnica DTW implică compararea unei combinații de două semnale, este necesar să se construiască o colecție de semnale de referință care să servească drept punct de referință pentru compararea tuturor celorlalte semnale.

Prin aplicarea metodei DTW între un semnal de referință și un alt semnal, vom obține matricea de deformare care conține coeficienții care pot fi utilizați pentru a comprima sau extinde un semnal pentru a obține o aliniere ideală între semnale.

Deoarece algoritmul DTW funcționează foarte bine atunci când se compară semnale relativ similare, ar fi optimă alegerea câtorva semnale de probă din fiecare test și stabilirea lor ca semnale de referință. Pentru procesul de selecție, propunem o primă reprezentare a semnalelor într-o formă bidimensională, în care pot fi grupate toate semnalele dintr-un test. În conformitate cu tehnica DTW, vom utiliza două semnale de referință aleatoare. Această proiecție a semnalelor în spațiul coeficienților rezultă din algoritmul DTW cu aceste semnale de referință aleatoare  $T = [T_1, T_2]$ . În urma executării grupării DBSCAN, va rezulta o determinare algoritmică a mai multor grupuri. Pentru a selecta semnalele de referință, este necesar să se stabilească o tehnică de extragere a centroizilor fiecărui cluster individual.

## 4.2.2 Construcția spațiului caracteristic al semnalelor analogice

Algoritmul bazat pe DTW va fi aplicat între semnalele de referință specificate în etapa anterioară și toate semnalele din cadrul unui test. Prin urmare, fiecărui semnal unic  $i$  se va atribui un set de coeficienți  $displaystyle C_j$  care definesc gradul său de similitudine cu semnalele de referință.

În prezent, semnalele obținute din fiecare test de verificare individual din baza de date posedă o reprezentare multidimensională, ceea ce permite analiza lor în scopul identificării oricărei separabilități liniare care ar putea exista între grupurile de semnale. Implementarea unei strategii de grupare în experimentele de laborator poate diferenția potențial semnalele care provin din experimente cu comportament așteptat de cele care rezultă din experimente de eșec sau comportamente anormale care prezintă valori aberante. Acest lucru poate fi facilitat prin standardizarea semnalelor de intrare ale circuitelor integrate analogice de testare. Spațiul caracteristic este construit cu ajutorul DBSCAN pentru a alege automat semnalele de referință. Această abordare este avantajoasă, deoarece urmărește minimizarea numărului de dimensiuni sau caracteristici implicate.

## 4.3 Model de rețea neuronală pentru semnale analogice

### 4.3.1 Prezentare generală

În această cercetare, oferim o metodă de reprezentare pentru semnalele analogice IC. Pasul inițial în procesul de reprezentare este selectarea caracteristicilor relevante din datele achiziționate. CNN-urile au fost selectate pentru a obține cele mai relevante

atribute în cel mai eficient mod. CNN a fost utilizat în mod eficient în diverse cazuri care necesită extragerea automată a caracteristicilor, cum ar fi [KHH18] și [RMX<sup>+</sup>19].

### 4.3.2 Model de rețea neuronală convoluțională

Obiectivul acestui model AE-CNN este de a extrage în mod eficient caracteristicile esențiale ale imaginilor de intrare prin comprimarea coeficienților rețelei în stratul intermediar. Autoencoderele sunt echipate cu un strat intermediar cunoscut în mod obișnuit sub numele de "gât de gâtul de sticlă", care are o dimensiune considerabil mai mică decât straturile de intrare și de ieșire.

Utilizarea rețelelor neuronale convoluționale este impusă de necesitatea de a prelua automat proprietățile semnalului pentru mai multe evenimente și tipuri de teste de circuite integrate analogice. În plus, acest lucru este crucial atunci când se tratează semnale care pot oferi apariții neașteptate pentru care este necesară o reprezentare robustă pentru a le evidenția pe parcursul procesului de vizualizare. Metodologiile convenționale necesită proiectarea manuală a descriptorilor pentru evenimente specifice, care este un proces plictisitor din cauza naturii sale empirice. Rezultatul nu este garantat a fi optim, deoarece necesită luarea în considerare a numeroase teste și circuite integrate analogice.

### 4.3.3 Optimizarea hiperparametrilor autoencoderului

Am utilizat optimizarea bayesiană [AR08] pentru a îmbunătăți parametrii modelului CNN-AE. Am împărțit baza de date în 70% de antrenament și 30% de testare pentru a îmbunătăți și evalua modelul cu cea mai mare eficiență și pentru a preveni supraajustarea. Prin modificarea numărului de coeficienți ai stratului de convoluție, am evaluat mai multe configurații pentru cazul nostru de utilizare specific. S-a constatat că următoarele setări ale hiperparametrului din tabelul 4.3 au dat cele mai bune rezultate de grupare după executarea procedurii de optimizare a modelului.

Table 4.3 Hiperparametrii modelului CNN-AE găsiți după optimizarea bayesiană [GDB<sup>+</sup>23]

Hyperparameter	Optimal Value
CNN - kernel	2
MaxPooling - pool size	2
MaxPooling - strides	2
Dropout - rate	0.3
Activation function	ReLU
Bottleneck - No. of Coefficients	128

## 4.4 Analiza metricilor de grupare

### 4.4.1 Overview

Modul în care evaluăm calitatea clusterizării este extrem de relevant, deoarece examinăm metodele de extragere a caracteristicilor prin prisma unei metrici de clusterizare. Selectarea metricii de clusterizare potrivite pentru cazul nostru de utilizare specific nu este o sarcină simplă, deoarece dorim să subliniem anumite comportamente observate pe parcursul procesului de verificare a circuitelor integrate analogice. Prin urmare, este necesară o examinare cuprinzătoare a diferitelor metrici pentru a identifica care este cea mai potrivită pentru aplicația noastră.

### 4.4.2 Sensitivity Analysis of Internal Evaluation Metrics

În acest studiu, am efectuat 13 teste cu distribuții gaussiene, iar rezultatele pot fi văzute în tabelul 4.4. În aceste teste, am ajustat distanța dintre clustere pentru testele 1-4, am creat unele clustere cu puncte adnotate incorect din testele 5-7 și am creat clustere cu valori aberante pentru testele 7-13. De asemenea,  $X_1$  și  $X_2$  sunt doi vectori aleatori care constituie baza pentru distribuțiile normale multivariate pentru fiecare set de puncte generate.

Table 4.4 Comparison between several clustering metrics

Id	No. of Points	Davies-Bouldin	Silhouette	CH	FM
Set 1	400	0.149	0.851	6500	1
Set 2	400	0.424	0.686	1527	0.985
Set 3	400	0.798	0.426	444	0.794
Set 4	400	11.4	0.026	2	0.506
Set 5	403	0.238	0.825	4051	0.985
Set 6	410	0.264	0.788	2438	0.952
Set 7	430	0.394	0.676	1155	0.870
Set 8	401	0.206	0.852	4449	1
Set 9	403	0.233	0.834	2351	1
Set 10	410	0.281	0.810	1251	1
Set 11	410	0.474	0.755	322	0.952
Set 12	410	0.464	0.759	333	0.952
Set 13	200	0.211	0.850	3186	1

## 4.5 Baza de date

În acest studiu, baza de date a fost extinsă de la zece teste constând în 2950 de semnale la treizeci de teste de semnale, totalizând 10200 de semnale. Fiecare test a fost compus

dintr-un număr limitat de categorii de semnale. Pentru fiecare test distinct de semnale, un specialist a grupat semnalele în două sau trei clase și a etichetat fiecare semnal cu clasa respectivă. Aceste etichete indică adevărul de bază pe care îl vom lua în considerare în timp ce evaluăm doar performanța de grupare. Deoarece folosim tehnici de învățare automată nesupravegheată, este recomandat să comparăm rezultatele cu etichetele pentru a afișa performanța globală.

## **4.6 Cazuri de utilizare experimentală a grupării semnalelor**

În acest capitol, am încercat să determinăm care dintre cei trei algoritmi propuși anterior este capabil să extragă cele mai bune caracteristici din semnalele analogice măsurate în timpul procesului de verificare. Pentru experimente, am utilizat o bază de date mai exhaustivă decât cea utilizată pentru a evalua eficacitatea algoritmului bazat pe SIFT în capitolul precedent. După cum se arată în tabelul 4.5, procedura bazată pe DTW produce cele mai bune rezultate pentru marea majoritate a testelor de semnal. Cazurile în care tehnica CNN-AE nu este depășită sunt cele care prezintă un grad ridicat de diferențiere a claselor și în care distanța dintre cluster este semnificativă.

## **4.7 Rezumat și concluzii**

În acest capitol, am prezentat și am demonstrat validitatea strategiilor de clusterizare pentru optimizarea procesului de verificare post-silicon folosind o metodă CNN-AE și un algoritm bazat pe DTW. Strategia CNN-AE a fost realizată prin combinarea rețelelor neuronale convoluționale și a rețelelor autocodificatoare capabile să genereze un spațiu de caracteristici adecvat pentru semnale analogice, în timp ce algoritmul bazat pe DTW a fost implementat prin utilizarea distanței de deformare a matricei DTW ca caracteristici cu referințe de formă de undă alese automat. Comparând metricile de puritate și Davies-Bouldin propuse în această lucrare cu abordările de grupare bazate pe SIFT dezvoltate anterior, am demonstrat că ambele metode din acest capitol sunt superioare. În ciuda performanțelor comparabile de evaluare externă, măsurate prin metricile de puritate, metoda de grupare bazată pe DTW prezintă un avantaj distinct față de omologul său atunci când este evaluată intern.

În acest capitol, am prezentat și validat două metode eficiente de grupare a semnalelor de răspuns la teste de circuit integrat pe baza similarității lor vizuale. Ca urmare a capacității algoritmilor propuși în acest capitol de a distinge între formele distincte de semnal, putem spune că avem un impact asupra metodologiei de verificare analogică post-silicon I prin reducerea cantității de efort manual necesar pentru identificarea semnalelor aberante.

Table 4.5 Comparație între rezultatele SIFT, CNN-AE și DTW

Test Name	Purity (%)			Davies-Bouldin		
	SIFT	CNN-AE	DTW	SIFT	CNN-AE	DTW
Test 1	100	100	100	0.365	0.281	<b>0.192</b>
Test 2	100	100	100	0.544	0.255	<b>0.179</b>
Test 3	100	100	100	<b>0.002</b>	0.007	0.004
Test 4	98	98	100	0.546	0.239	<b>0.152</b>
Test 5	98	100	100	0.976	0.334	<b>0.188</b>
Test 6	97	100	100	1.422	0.318	<b>0.001</b>
Test 7	92	96	96	0.681	0.552	<b>0.466</b>
Test 8	93	97	100	0.421	0.418	<b>0.211</b>
Test 9	100	100	100	0.420	0.277	<b>0.009</b>
Test 10	88	99	100	1.526	0.401	<b>0.250</b>
Test 11	100	100	100	0.592	0.363	<b>0.176</b>
Test 12	100	100	100	0.370	0.229	<b>0.194</b>
Test 13	72.2	91	100	1.423	0.729	<b>0.015</b>
Test 14	73.6	77.6	100	1.488	1.198	<b>0.265</b>
Test 15	100	100	100	0.470	0.372	<b>0.197</b>
Test 16	100	100	100	0.311	0.214	<b>0.163</b>
Test 17	100	100	100	0.249	<b>0.152</b>	0.171
Test 18	94.3	100	100	0.527	0.418	<b>0.364</b>
Test 19	64.6	59.6	90.6	1.422	3.017	<b>0.728</b>
Test 20	94.8	96.8	100	0.663	0.558	<b>0.242</b>
Test 21	96.3	100	100	0.486	0.233	<b>0.219</b>
Test 22	100	100	100	0.336	0.185	<b>0.091</b>
Test 23	100	100	100	0.103	<b>0.069</b>	0.085
Test 24	87.3	100	100	0.881	0.365	<b>0.222</b>
Test 25	98.6	100	100	0.647	0.140	<b>0.064</b>
Test 26	100	100	100	0.335	0.277	<b>0.255</b>
Test 27	99.3	100	100	0.481	0.368	<b>0.299</b>
Test 28	100	100	100	0.164	<b>0.082</b>	0.104
Test 29	62	64	88	0.929	1.049	<b>0.588</b>
Test 30	74	86	92	0.953	0.729	<b>0.662</b>

# Chapter 5

## Asistarea verificării fabricației de wafere

### 5.1 Prezentare generală

Această lucrare abordează problema diminuării cantității de muncă umană necesară pentru analiza senzorilor de producție. Acest lucru este realizat cu ajutorul algoritmului bazat pe DTW și a analizei de grupare metrică în combinație cu un clasificator SVM sau cu metrica Davies-Bouldin. Utilizarea algoritmului bazat pe DTW este motivată de eficacitatea sa în codificarea disimilarităților dintre semnale analogice, care este exemplificată în software-ul de analiză a semnalelor vocale [HPB19] și, de asemenea, în lucrarea noastră [GDB<sup>+</sup>22]. Clusterizarea este o metodologie eficientă care poate sprijini fabricarea de circuite integrate analogice. Procesul de grupare a datelor poate ajuta la identificarea modelelor fundamentale și a modurilor de defecțiune, facilitând astfel optimizarea procesului de fabricație și sporind randamentul.

### 5.2 Baza de date

Setul de date de validare pentru abordarea propusă include 972 de plachete măsurate și etichetate (171/801 eșecuri/ eșecuri reușite). Fiecare plachetă este definită de 56 de forme de undă ale senzorilor obținute în timpul metodei de testare, totalizând 54 432 de forme de undă distincte. În plus, este prezentată clasificarea fiecăruia dintre cele 56 de teste aferente. Expertii au generat aceste adnotări și clasificări ale setului de date utilizând o tehnică de examinare vizuală. Datorită faptului că studiul a cuprins un total de 54.432 de forme de undă unice, putem afirma cu încredere că setul de date utilizat este suficient de mare pentru a verifica metodologia propusă. Acest set de date nu a putut fi utilizat pentru scenariul de grupare descris în capitolele 3 și 4, deoarece etichetele au fost atribuite diferit. Acest lucru a fost determinat nu de forma de undă a semnalelor, ca în cazurile anterioare, ci de comportamentul plachetei.

### 5.3 Detectarea eșecurilor la wafere

Tehnica de extragere a caracteristicilor care ar putea evidenția (di)similitudinile formelor de undă se bazează pe Dynamic Time Warping și măsurile sale conexe, așa cum sunt descrise în [HPB19], [MCC<sup>+</sup>15]. Fiecare set de forme de undă ale senzorilor prezintă un anumit grad de similaritate în raport cu plachetele care nu prezintă defecte, abordarea bazată pe DTW a fost aleasă ca algoritm de extragere a caracteristicilor. Raționamentul din spatele alegerii algoritmului bazat pe DTW este atribuit performanței sale superioare în verificarea circuitelor integrate, în comparație cu algoritmi alternativi. De asemenea, rezistența algoritmului la zgomot și invarianța la lungime au fost factori importanți.

Am utilizat un clasificator SVM antrenat pentru a măsura gradul de separare între fiecare formă de undă a senzorului, prin localizarea hiperplanului de separare între eșantioanele defecte și cele care nu sunt defecte. Gradul de separabilitate a claselor, determinat de precizia clasificării SVM, este un parametru de clasificare a senzorilor în ceea ce privește corelația cu plachetele defecte. Un parametru alternativ de clasificare este o metrică de cluster Davies-Bouldin aplicată în spațiul caracteristic.

Pentru determinarea separabilității celor două clase în interiorul spațiului caracteristic, am utilizat precizia de clasificare a unui clasificator neliniar. Am ales un nucleu gaussian pentru clasificatorul SVM [? ], datorită caracteristicii sale de marjă maximă de clasificare optimă [ZG08]. Am antrenat un SVM pentru fiecare senzor de testare, cu 75% din date utilizate pentru antrenament și 25% pentru testare. Abordarea de instruire a implicat optimizarea hiperparametrilor scalei nucleului gaussian și a constrângerii box, utilizând un algoritm de optimizare bayesiană ca funcție de achiziție [FIG18]. Am utilizat, de asemenea, o standardizare echilibrată a ponderilor pentru a compensa natura dezechilibrată a setului de date.

Acuratețea clasificării pe setul de testare a fost utilizată ca o măsură de clasificare pentru fiecare set de caracteristici ale formei de undă a senzorului. Factorii principali care contribuie la eșec pot fi atribuiți testelor care oferă cea mai mare precizie în clasificarea proprietăților formei de undă. Metrica Davies-Bouldin a fost utilizată pentru a evalua gradul de distinctivitate dintre cele două categorii în spațiul caracteristic în scopul comparației. Hiperparametrii suplimentari ai SVM cuprind parametrul de grad și parametrul de coeficient, care se aplică exclusiv la nucleele polinomiale. În plus, dimensiunea cache-ului și criteriile de convergență sunt utilizate pentru a regla consumul de memorie în timpul antrenamentului și, respectiv, criteriul de oprire a algoritmului de optimizare.

Selectarea hiperparametrilor adecvați pentru SVM depinde de problema specifică în cauză și de atributele inerente ale datelor. Experimentarea și validarea, care ajută la alegerea valorilor adecvate ale hiperparametrilor, pot contribui la obținerea unei performanțe optime pentru o anumită problemă.



## 5.4 Rezultate experimentale

În timpul acestui experiment, experții au evaluat manual care dintre senzorii existenți în setul de date au legătură cu comportamentul de defecțiune și au determinat un ordin de semnificație în ceea ce privește acest comportament. Prin urmare, în tabelul 5.1 am ordonat primii 20 de senzori așa cum au fost determinați manual de către experți. Această clasificare manuală va servi drept bază pentru analiza rezultatelor. Este important să recunoaștem că clasificarea poate fi o sarcină care necesită mult timp, putând dura câteva săptămâni. Prin urmare, eforturile noastre s-au concentrat pe găsirea unor metode de a aranja automat acești senzori aproximativ în aceeași ordine, pentru a economisi timp.

Deși cele trei metodologii de clasificare prezintă variații, ele mențin o ordine de clasificare similară în comparație cu referința de clasificare manuală. Toate cele trei metode de clasificare produc o listă identică a celor mai bine clasificați senzori. Constatările indică faptul că sistemul are capacitatea de a automatiza evaluarea obligatorie a ratingurilor senzorilor în ceea ce privește corelația acestora cu o defecțiune repetitivă, un proces care implică de obicei expertiză umană și un efort manual semnificativ, menținând în același timp un nivel similar de acuratețe.

Table 5.1 Clasamentul primilor 20 de senzori pe baza preciziei SVM și a metricii Davies-Bouldin

Manual Ranking	SVM Accuracy [%]	Davies-Bouldin Score
Sensor 12	67.07	0.88
Sensor 4	66.66	0.96
Sensor 35	65.02	0.94
Sensor 7	62.69	0.98
Sensor 5	64.60	1.05
Sensor 50	59.25	1.26
Sensor 35	60.08	1.14
Sensor 21	61.31	1.51
Sensor 17	58.02	1.36
Sensor 38	57.01	1.51
Sensor 28	57.61	1.45
Sensor 16	56.91	1.63
Sensor 3	54.22	1.52
Sensor 36	50.11	1.55
Sensor 41	51.55	1.91
Sensor 29	53.39	1.98
Sensor 44	53.61	1.92
Sensor 19	52.18	1.98
Sensor 31	49.36	1.89
Sensor 14	47.88	2.03

În cea de-a treia coloană a tabelului 5.1, am calculat indicele Davies-Bouldin pentru fiecare senzor, care este o metodă mai simplă și mai rapidă de determinare a separabilității liniare. Atunci când se ia în considerare acest indice, este important să se țină cont de faptul că o valoare mai mică corespunde unui grad mai mare de separabilitate. După cum se arată în tabelul 5.1, acest parametru este validat de ordinea stabilită manual.

## 5.5 Concluzii

O metodă automată de clasificare a senzorilor asociați cu defecțiuni este o abordare viabilă pentru a sprijini verificarea fabricației de wafere. Procesul presupune identificarea senzorilor care sunt potențial legați de defecțiune și, ulterior, ierarhizarea acestora pe baza semnificației sau a relevanței lor pentru defecțiune. Pentru a executa această metodologie, poate fi necesară colectarea de informații de la senzori și examinarea acestora pentru a recunoaște tiparele sau tendințele care ar putea fi asociate cu defecțiunea. O abordare potențială constă în utilizarea analizei statistice sau a metodologiilor de învățare automată pentru a stabili asocieri între datele senzorilor și apariția defecțiunilor sistemului.

După clasificarea senzorilor, este plauzibil să se utilizeze aceste date pentru a stabili motivul fundamental al defecțiunii. În cazul în care un anumit senzor se dovedește a fi clasat în mod constant ca fiind cel mai important senzor, este plauzibil ca acesta să servească drept factor de bază pentru defecțiune. În mod alternativ, poate fi necesar să se efectueze cercetări suplimentare asupra senzorilor cu prioritate ridicată pentru a determina cauza fundamentală. Metodologia utilizată în abordarea noastră produce rezultate echivalente cu cele obținute prin aplicarea expertizei umane. Cu toate acestea, se consideră că este mai eficientă, în special în ceea ce privește coerența procedurii de clasificare și reducerea efortului manual.

Prin utilizarea fie a acurateții de clasificare SVM, fie a metricii Davies-Bouldin asupra caracteristicilor recuperate de metoda bazată pe DTW, am obținut clasificări ale senzorilor care sunt comparabile cu cele obținute de experții umani. Pe baza eficacității analizei senzorilor, se poate concluziona că există o influență semnificativă în ceea ce privește sprijinirea fabricării de circuite integrate.

# Chapter 6

## Concluzii generale

Scopul acestui studiu a fost de a investiga utilizarea metodelor de învățare automată pentru optimizarea verificării circuitelor integrate analogice. Necesitatea acestor tehnici apare din cauza cantității considerabile de muncă umană necesară pentru verificarea circuitelor integrate analogice, care poate fi redusă substanțial.

### 6.1 Obiective generale și rezultate

Prin combinarea semnalelor individuale în clustere mai mari, primele două faze ale acestui proces sunt capabile să atenueze o parte din provocarea de a îmbunătăți eficiența cu care sunt verificate circuitele integrate. Acest lucru este deosebit de util pentru procesul de verificare, deoarece înseamnă că inginerii trebuie să evalueze vizual doar un număr limitat de clustere ale căror semnale sunt în principal identice, spre deosebire de examinarea vizuală a mii de semnale. În cea de-a treia etapă, am îmbunătățit procesul de identificare a cauzei principale a unei probleme, spre deosebire de simpla recunoaștere a existenței acesteia, ridicând astfel abordarea noastră la un nivel superior.

#### **O1. Reducerea timpului necesar pentru verificarea circuitelor integrate analogice**

O cantitate semnificativă de timp este alocată procesului de verificare în dezvoltarea circuitelor integrate, care a devenit foarte costisitor pe măsură ce cerințele au crescut. Prin urmare, este necesar să se reducă munca manuală implicată în acest proces și să se accelereze verificarea circuitelor integrate.

##### **O1.1. Algoritm de extragere a caracteristicilor pentru comportamentele așteptate**

Pentru a atinge acest obiectiv, am dezvoltat un algoritm capabil să extragă în mod optim caracteristicile anumitor comportamente care apar în mod regulat în timpul verificării circuitelor integrate. Acest subiect a fost discutat pe larg în capitolul 3, secțiunea 3.2.

##### **O1.2. Metoda de optimizare a invarianței semnalului pentru condiții de simulare**

Acest obiectiv este atins prin implementarea unei compresii de tip DCT și a unei abordări Bag-Of-Words. Acest obiectiv este descris mai detaliat în capitolul 3, secțiunea 3.2.3.

### **O1.3. Gruparea semnalelor care conțin evenimente similare**

Rezultatul acestui obiectiv a fost reprezentat de algoritmi de grupare aplicați spațiului multidimensional al caracteristicilor în scopul grupării semnalelor similare. Această etapă a fost descrisă în capitolul 3, secțiunea 3.2.5.

## **O2. Permitearea unei acoperiri sporite și mai rapide a procesului de verificare a circuitelor integrate**

Combi-națiile de parametri trebuie verificate pentru a asigura o acoperire excelentă și un standard de calitate ridicat, ceea ce reprezintă o provocare pentru procesul de verificare pe măsură ce numărul de cerințe crește.

### **O2.1. Metode eficiente de extragere a caracteristicilor pentru comportamente neobișnuite**

Acest obiectiv a fost atins prin dezvoltarea a 2 algoritmi de învățare automată capabili să extragă caracteristici utile din orice tip de formă de undă cu scopul de a separa comportamente neobișnuite care pot fi izolate, în capitolul 4, secțiunea 4.2 și secțiunea 4.3.

### **O2.2. Definirea celor mai potrivite metri-ci de grupare pentru semnalele analogice IC**

Acest obiectiv este atins prin efectuarea unui studiu asupra mai multor metri-ci de grupare pentru a selecta cea mai bună metrică adecvată pentru aplicațiile de verificare a circuitelor integrate, în capitolul 4, secțiunea 4.4

### **O2.3. Metoda de optimizare a rezistenței la zgomot pentru condiții de laborator**

Acest obiectiv este atins prin implementarea unui proces de selecție DB-SCAN în scopul eliminării atât a informațiilor redundante, cât și a zgomotului legat de măsurare în capitolul 4, secțiunea 4.2.

## **O3. Optimizarea verificării producției de wafere**

În timpul testării circuitelor post-siliciu, este vital să se determine sursa comportamentelor neașteptate repetate pentru a asigura o fabricație optimă și fără erori.

### **O2.1. Îmbunătățirea metodologiei de testare a liniilor de producție**

Acest obiectiv a fost atins prin adaptarea celei mai eficiente metode de învățare automată pentru extragerea atributelor din cele pe care le-am evaluat și prin aplicarea acestora la o bază de date diferită de plachete. Acest lucru a fost prezentat în capitolul 5.

## **O2.2. Abordare autonomă de clasificare a senzorilor pentru a sprijini fabricarea IC**

Acest lucru este realizat prin implementarea unei metode automate unei metode o clasificare a senzorilor pe baza corelației acestora cu un comportament repetitiv de defecțiune. Acest lucru a fost prezentat în capitolul 5 și 5.4.

## **6.2 Contribuții originale**

Principalele contribuții aduse de Capitolul 3:

- Am creat un descriptor inovator pentru această metodă, bazat pe o tehnică de viziune computerizată care recunoaște punctele de interes [GND<sup>+</sup>20].
- Furnizarea unei metode de vizualizare a spațiului caracteristic pentru a optimiza faza de verificare în care este necesară supravegherea umană [GND<sup>+</sup>20].

Principale aduse de capitolul 4:

- Dezvoltarea unei tehnici de extragere a caracteristicilor nesupravegheate care se bazează pe deformarea dinamică a timpului și este capabilă să caracterizeze cu precizie atât evenimentele tipice, cât și cele neobișnuite. [GDB<sup>+</sup>22].
- Elaborarea unui algoritm bazat pe o rețea neuronală de tip autocodificator care este optimizat pentru semnalele de interes și este capabil să învețe anumite caracteristici și să producă o reprezentare a semnalelor cu foarte puțini coeficienți. [GDB<sup>+</sup>23].

Principalele contribuții aduse de Capitolul 5:

- Dezvoltarea unei tehnici de extragere a caracteristicilor prin învățare automată pentru a permite utilizarea acestora în contextul metodologiei de testare a liniilor de producție [GBD<sup>+</sup>22].
- Elaborarea unei tehnici automate prin îmbinarea metodelor de clasificare cu metricile constatărilor de grupare a senzorilor. În consecință, obținerea unei clasificări a senzorilor pe baza asocierii lor cu un comportament de defecțiune recurent. [GBD<sup>+</sup>22].

## **6.3 Lista publicațiilor originale**

1. [GND<sup>+</sup>20] Andrei Gaita, Georgian Nicolae, Emilian C. David, Andi Buzo, Corneliu Burileanu, and Georg Pelz. A sift-based waveform clustering method for aiding analog/mixed-signal ic verification. In *2020 IEEE European Test Symposium (ETS)*, pages 1–2, 2020, ISI WOS:000615974000037

2. [GDB<sup>+</sup>22] A. Gaita, E. David, A. Buzo, H. Cucu, and G. Pelz. Waveform clustering based on dynamic time warping used in analog ic verification. In *2022 International Symposium ELMAR*, pages 49–52, 2022, ISI WOS:000935062500011
3. [GBD<sup>+</sup>22] A. Gaita, A. Buzo, E. David, H. Cucu, and G. Pelz. A machine learning based wafer test ranking for root cause analysis. In *2022 International Symposium ELMAR*, pages 45–48, 2022, ISI WOS:000935062500010
4. [GDB<sup>+</sup>23] A. Gaita, E. David, A. Buzo, M. Grigore, C. Burileanu, H. Cucu, and G. Pelz. Convolutional neural network model used for aiding ic analog/mixed signal verification. *UNIVERSITY POLITEHNICA OF BUCHAREST SCIENTIFIC BULLETIN SERIES C-ELECTRICAL ENGINEERING AND COMPUTER SCIENCE*, 85(2):151–162, 2023, ISI WOS:001015488500009

## 6.4 Lista rapoartelor tehnice

1. [Gai19b] A. Gaita. Ic optimization methods based on machine learning. *Technical Report No. 1, University Politehnica of Bucharest*, June 2019
2. [Gai19a] A. Gaita. Augmented ic analog signals verification with waveform clustering. *Technical Report No. 2, University Politehnica of Bucharest*, December 2019
3. [Gai20] A. Gaita. Assisted analog/mixed-signal integrated circuit verification using a dtw-based waveform clustering. *Technical Report No. 3, University Politehnica of Bucharest*, June 2020

## 6.5 Lucrări viitoare

Domeniile potențiale de cercetare sau îmbunătățiri ale activității actuale care ar putea îmbunătăți validarea circuitelor integrate analogice includ următoarele subiecte:

- O extindere notabilă a bazelor de date care cuprind semnale adnotate, destinate în special aplicării metodologiilor de învățare aprofundată care necesită cantități substanțiale de date.
- Îmbunătățirea tehnicilor existente de vizualizare a spațiului caracteristic pentru a evidenția atributele specifice pentru diverse cazuri de utilizare care pot fi semnificative.
- Integrarea algoritmului de determinare a cauzelor profunde în cadrul de verificare, permițând identificarea automată a senzorilor cu cea mai mare corelație cu defecțiunile repetitive.

# References

- [ANR74] N. Ahmed, T. Natarajan, and K.R. Rao. Discrete cosine transform. *IEEE Transactions on Computers*, C-23(1):90–93, 1974.
- [AR08] Chang Wook Ahn and R. S. Ramakrishna. On the scalability of real-coded bayesian optimization algorithm. *IEEE Transactions on Evolutionary Computation*, 12(3):307–322, 2008.
- [BY21] Adil Abdu Bushra and Gangman Yi. Comparative analysis review of pioneering dbscan and successive density-based clustering algorithms. *IEEE Access*, 9:87918–87935, 2021.
- [CH74] T. Caliński and J Harabasz. A dendrite method for cluster analysis. *Communications in Statistics*, 3(1):1–27, 1974.
- [CK07] Henry Chang and Ken Kundert. Verification of complex analog and rf ic designs. *Proceedings of the IEEE*, 95(3):622–639, 2007.
- [CTGK22] Animesh Basak Chowdhury, Benjamin Tan, Siddharth Garg, and Ramesh Karri. Robust deep learning for ic test problems. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 41(1):183–195, 2022.
- [Den20] Dingsheng Deng. Dbscan clustering algorithm based on density. In *2020 7th International Forum on Electrical Engineering and Automation (IFEAA)*, pages 949–953, 2020.
- [dSCF<sup>+</sup>12] Marcilio C.P. de Souto, André L.V. Coelho, Katti Faceli, Tiemi C. Sakata, Viviane Bonadia, and Ivan G. Costa. A comparison of external clustering evaluation indices in the context of imbalanced data sets. In *2012 Brazilian Symposium on Neural Networks*, pages 49–54, 2012.
- [DYL08] Hong Duan, Ruohe Yan, and Kunhui Lin. Research on face recognition based on pca. In *2008 International Seminar on Future Information Technology and Management Engineering*, pages 29–32, 2008.
- [FIPG18] Andrew Christian Flores, Rogelyn I. Icoy, Christine F. Peña, and Ken D. Gorro. An evaluation of svm and naive bayes with smote on sentiment analysis data set. In *2018 International Conference on Engineering, Applied Sciences, and Technology (ICEAST)*, pages 1–4, 2018.
- [FM83] E. B. Fowlkes and C. L. Mallows. A method for comparing two hierarchical clusterings. *Journal of the American Statistical Association*, 78(383):553–569, 1983.
- [Gai19a] A. Gaita. Augmented ic analog signals verification with waveform clustering. *Technical Report No. 2, University Politehnica of Bucharest*, December 2019.

## References

- [Gai19b] A. Gaita. Ic optimization methods based on machine learning. *Technical Report No. 1, University Politehnica of Bucharest*, June 2019.
- [Gai20] A. Gaita. Assisted analog/mixed-signal integrated circuit verification using a dtw-based waveform clustering. *Technical Report No. 3, University Politehnica of Bucharest*, June 2020.
- [GBD<sup>+</sup>22] A. Gaita, A. Buzo, E. David, H. Cucu, and G. Pelz. A machine learning based wafer test ranking for root cause analysis. In *2022 International Symposium ELMAR*, pages 45–48, 2022.
- [GDB<sup>+</sup>22] A. Gaita, E. David, A. Buzo, H. Cucu, and G. Pelz. Waveform clustering based on dynamic time warping used in analog ic verification. In *2022 International Symposium ELMAR*, pages 49–52, 2022.
- [GDB<sup>+</sup>23] A. Gaita, E. David, A. Buzo, M. Grigore, C. Burileanu, H. Cucu, and G. Pelz. Convolutional neural network model used for aiding ic analog/mixed signal verification. *UNIVERSITY POLITEHNICA OF BUCHAREST SCIENTIFIC BULLETIN SERIES C-ELECTRICAL ENGINEERING AND COMPUTER SCIENCE*, 85(2):151–162, 2023.
- [GK12] Bernhard C. Geiger and Gernot Kubin. Relative information loss in the pca. In *2012 IEEE Information Theory Workshop*, pages 562–566, 2012.
- [GND<sup>+</sup>20] Andrei Gaita, Georgian Nicolae, Emilian C. David, Andi Buzo, Corneliu Burileanu, and Georg Pelz. A sift-based waveform clustering method for aiding analog/mixed-signal ic verification. In *2020 IEEE European Test Symposium (ETS)*, pages 1–2, 2020.
- [GS19] Guruprasad and Kumara Shama. Design and verification of analog integrated circuits using free or open source eda tools. In *2019 International Conference on Communication and Electronics Systems (ICCES)*, pages 1–6, 2019.
- [GXGM19] Georges Gielen, Nektar Xama, Karthik Ganesan, and Subhasish Mitra. Review of methodologies for pre- and post-silicon analog verification in mixed-signal socs. In *2019 Design, Automation and Test in Europe Conference and Exhibition (DATE)*, pages 1006–1009, 2019.
- [HPB19] Jae Yeol Hong, Seung Hwan Park, and Jun-Geol Baek. Segmented dynamic time warping based signal pattern classification. In *2019 IEEE International Conference on Computational Science and Engineering (CSE) and IEEE International Conference on Embedded and Ubiquitous Computing (EUC)*, pages 263–265, 2019.
- [ISAS21] Khairul Nurrazianna Ismail, Ali Seman, and Khyrina Airin Fariza Abu Samah. A comparison between external and internal cluster validity indices. In *2021 IEEE 11th International Conference on System Engineering and Technology (ICSET)*, pages 229–233, 2021.
- [JJJ<sup>+</sup>20] Chen Jie, Zhang Jiyue, Wu Junhui, Wu Yusheng, Si Huiping, and Lin Kaiyan. Review on the research of k-means clustering algorithm in big data. In *2020 IEEE 3rd International Conference on Electronics and Communication Engineering (ICECE)*, pages 107–111, 2020.



- [KC06] Ken Kundert and Henry Chang. Verification of complex analog integrated circuits. In *IEEE Custom Integrated Circuits Conference 2006*, pages 177–184, 2006.
- [KD16] Ergin Kılıç and Erdi Doğan. Real-time feature extraction from emg signals. In *2016 24th Signal Processing and Communication Application Conference (SIU)*, pages 113–116, 2016.
- [KHH18] Shoji Kido, Yasusi Hirano, and Noriaki Hashimoto. Detection and classification of lung abnormalities by use of convolutional neural network (cnn) and regions with cnn features (r-cnn). In *2018 International Workshop on Advanced Image Technology (IWAIT)*, pages 1–4, 2018.
- [KMH17] Ichwanul Muslim Karo Karo, Kiki MaulanaAdhinugraha, and Arief Fatchul Huda. A cluster validity for spatial clustering based on davies bouldin index and polygon dissimilarity function. In *2017 Second International Conference on Informatics and Computing (ICIC)*, pages 1–6, 2017.
- [KuRA<sup>+</sup>14] Kamran Khan, Saif ur Rehman, Kamran Aziz, Simon James Fong, Sababady Sarasvady, and Amrita Vishwa. Dbscan: Past, present and future. *The Fifth International Conference on the Applications of Digital Information and Web Technologies (ICADIWT 2014)*, pages 232–238, 2014.
- [KYY20] Naveen Kumar, Sanjay Kumar Yadav, and Divakar Singh Yadav. Similarity measure approaches applied in text document clustering for information retrieval. In *2020 Sixth International Conference on Parallel, Distributed and Grid Computing (PDGC)*, pages 88–92, 2020.
- [LLX<sup>+</sup>10] Yanchi Liu, Zhongmou Li, Hui Xiong, Xuedong Gao, and Junjie Wu. Understanding of internal clustering validation measures. In *2010 IEEE International Conference on Data Mining*, pages 911–916, 2010.
- [LLX<sup>+</sup>13] Yanchi Liu, Zhongmou Li, Hui Xiong, Xuedong Gao, Junjie Wu, and Sen Wu. Understanding and enhancement of internal clustering validation measures. *IEEE Transactions on Cybernetics*, 43(3):982–994, 2013.
- [LO12] Jonathan Lilly and Sofia Olhede. Generalized morse wavelets as a superfamily of analytic wavelets. *IEEE Transactions on Signal Processing*, 60, 03 2012.
- [Low04] David G. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2):91–110, Nov 2004.
- [MCC<sup>+</sup>15] Victor Maus, Gilberto Câmara, Ricardo Cartaxo, Fernando M. Ramos, Alber Sanchez, and Gilberto Q. Ribeiro. Open boundary dynamic time warping for satellite image time series classification. In *2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pages 3349–3352, 2015.
- [MCC<sup>+</sup>22] Cristian Manolache, Alexandru Caranica, Horia Cucu, Andi Buzo, Cristian Diaconu, and Georg Pelz. Enhanced candidate selection algorithm for analog circuit verification. In *2022 International Semiconductor Conference (CAS)*, pages 137–140, 2022.

## References

- [MGK<sup>+</sup>05] H. Morgenstern, G. Groos, H. Kohne, M. Stecher, W. John, and H. Reichl. Algorithm for the automatic verification of complex mixed-signal ics regarding esd-stress. In *Research in Microelectronics and Electronics, 2005 PhD*, volume 1, pages 213–216 vol.1, 2005.
- [MR93] Andrzej Maćkiewicz and Waldemar Ratajczak. Principal components analysis (pca). *Computers and Geosciences*, 19(3):303–342, 1993.
- [PSJ15] Sakshi Patel, Shivani Sihmar, and Aman Jatain. A study of hierarchical clustering algorithms. In *2015 2nd International Conference on Computing for Sustainable Global Development (INDIACom)*, pages 537–541, 2015.
- [RHW<sup>+</sup>16] Weijie Ren, Min Han, Jun Wang, Dan Wang, and Tieshan Li. Efficient feature extraction framework for eeg signals classification. In *2016 Seventh International Conference on Intelligent Control and Information Processing (ICICIP)*, pages 167–172, 2016.
- [RJ18] K.V.S.N. Rama Rao and B. Manjula Josephine. Exploring the impact of optimal clusters on cluster purity. In *2018 3rd International Conference on Communication and Electronics Systems (ICCES)*, pages 754–757, 2018.
- [RMX<sup>+</sup>19] Anushree Ramanath, Saipreethi Muthusrinivasan, Yiqun Xie, Shashi Shekhar, and Bharathkumar Ramachandra. Ndzi versus cnn features in deep learning for land cover clasification of aerial images. In *IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium*, pages 6483–6486, 2019.
- [RS01] Mark Robertson and Robert Stevenson. Dct quantization noise in compressed images. 1, 01 2001.
- [RZWD15] Narender Rana, Yunlin Zhang, Donald Wall, and Bachir Dirahoui. Predictive data analytics and machine learning enabling metrology and process control for advanced node ic fabrication. In *2015 26th Annual SEMI Advanced Semiconductor Manufacturing Conference (ASMC)*, pages 313–319, 2015.
- [SMMS20] Akhilesh Kumar Singh, Shantanu Mittal, Prashant Malhotra, and Yash Vardhan Srivastava. Clustering evaluation by davies-bouldin index(dbi) in cereal data using k-means. In *2020 Fourth International Conference on Computing Methodologies and Communication (ICCMC)*, pages 306–310, 2020.
- [SN20] Ketan Rajshekhhar Shahapure and Charles Nicholas. Cluster quality analysis using silhouette score. In *2020 IEEE 7th International Conference on Data Science and Advanced Analytics (DSAA)*, pages 747–748, 2020.
- [VK16] Ankit Vij and Padmavati Khandnor. Validity of internal cluster indices. In *2016 International Conference on Computation System and Information Technology for Sustainable Solutions (CSITSS)*, pages 388–395, 2016.
- [XW10] Rui Xu and Donald C. Wunsch. Clustering algorithms in biomedical research: A review. *IEEE Reviews in Biomedical Engineering*, 3:120–154, 2010.
- [YWCW21] Yueyi Yang, Lide Wang, Huang Chen, and Chong Wang. An end-to-end denoising autoencoder-based deep neural network approach for fault diagnosis of analog circuit. *Analog Integr. Circuits Signal Process.*, 107(3):605–616, June 2021.

- [ZG08] Cai Zhili and Jiang Guiyan. Application of multiple svm classifier fusion technique in freeway automatic incident detection. In *2008 27th Chinese Control Conference*, pages 581–585, 2008.

