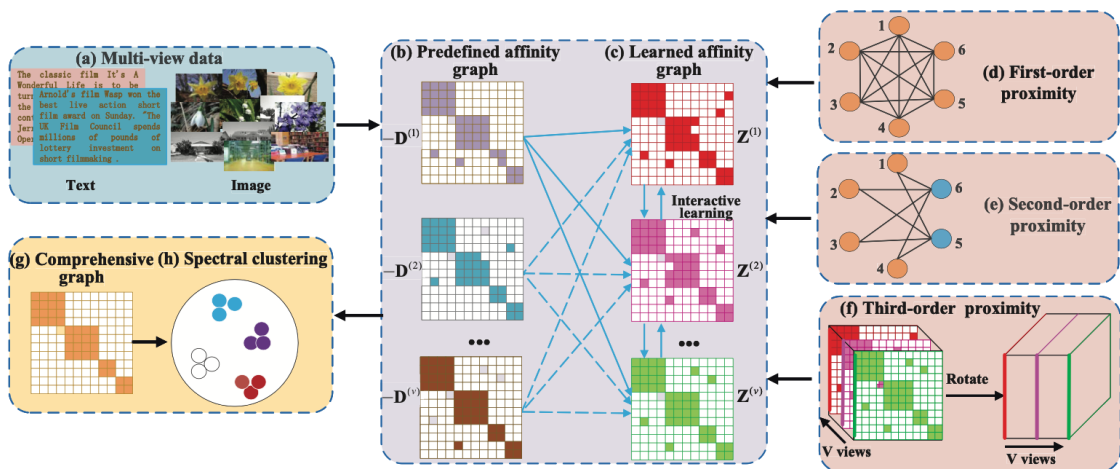


ISSUE: MAY 2023

# CTSOC-NCT NEWS ON CONSUMER TECHNOLOGY



The overview of the proposed MCGC

---

|           |                  |
|-----------|------------------|
| <b>2</b>  | EDITOR'S NOTE    |
| <b>3</b>  | COVER STORY      |
| <b>4</b>  | FEATURED PEOPLE  |
| <b>13</b> | FEATURED ARTICLE |

# TABLE OF CONTENTS

# EDITOR'S NOTE

On behalf of the Editorial Board of IEEE CTSoc News on Consumer Technology (NCT) editor-in-chief Wen-Huang Cheng and editors, Yafei Hou, Chuan-Ju Wang, Jianlong Fu and Loh Yuen Peng, I am happy to introduce the May issue of the NCT in 2023.

This issue starts with a cover story about the article Multi-view Comprehensive Graph Clustering published in the CTSoc's journal, IEEE Transactions on Consumer Electronics. In this paper, a new multi-view comprehensive graph clustering method is devised, which can thoroughly learn the similarity based on (1) first-order proximity (i.e., the direct relation of pairwise samples); (2) second-order proximity (i.e., the indirect relation of pairwise samples); and (3) third-order proximity (i.e., the three-order relation of multiple views).

Afterwards, the feature people provide an interview with Dr. Ghee Leng, Ooi who is a Co-Founder, AkiraKan. AkiraKan is a dynamic and innovative company specializing in geospatial intelligence, intelligent transportation systems, and weather technologies. With a strong focus on leveraging edge compute, advanced algorithms and AI technology, they provide cutting-edge products and services that empower organizations across various industries.

Finally, this issue presents a featured article brought by Massimo Martini, Riccardo Rosati and Luca Romeo, entitled Data augmentation strategy for generating realistic samples on defect segmentation task. In the paper a data-level solution is introduced to improve the generalization performance of semantic segmentation of surface defects. Specifically, the proposed approach includes a generative phase to simulate synthetic defects and a validation phase to verify that the synthetic image is as close as possible to the real one. With real experiments on a benchmark dataset, we demonstrated the effectiveness of our approach in a real-use scenario compared to other widely used data augmentation approaches for semantic segmentation for defect detection. The data augmentation approach allows minority classes to be balanced while improving overall generalization performance.

We hope you can enjoy your reading!

Luca Romeo  
Editor of NCT



---

**ARTICLE TITLE**

Multi-view Comprehensive Graph Clustering

**AUTHOR(S)**

Yanying Mei, Zhenwen Ren, Bin Wu, Tao Yang, Yanhua Shao

**JOURNAL TITLE**

IEEE TRANSACTIONS ON CONSUMER ELECTRONICS

**JOURNAL VOLUME AND ISSUE**

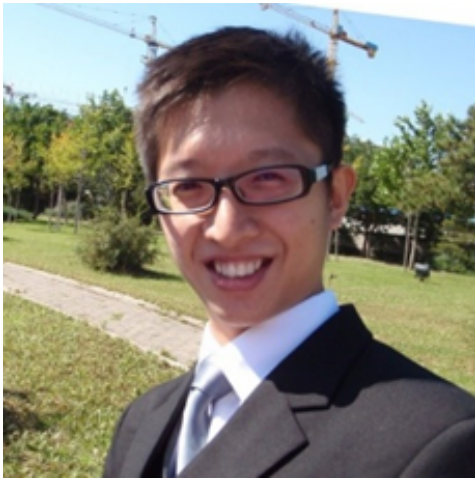
Early Access

**DATE OF THE ARTICLE**10 March 2023

---

Multi-view clustering algorithms have been successfully used in consumer electronic products, such as common digital cameras and unmanned vehicles. Currently, existing multi-view graph clustering (MVGC) methods learn the similarity of directly connected samples for clustering. However, these MVGC methods cannot fully consider the indirect relation among samples and high-order relation across multi-view data. In this paper, a new multi-view comprehensive graph clustering (MCGC) method is devised, which can thoroughly learn the similarity based on (1) first-order proximity (FOP) (i.e., the direct relation of pairwise samples); (2) second-order proximity (SOP) (i.e., the indirect relation of pairwise samples); and (3) third-order proximity (TOP) (i.e., the three-order relation of multiple views). Since the operations of these three components are iteratively carried out, the interaction between similarity learning can be encouraged, and the comprehensive graph can be generated effectively for clustering. In-depth experiments on six commonly benchmark datasets show the superiority of the MCGC method.

# FEATURED INTERVIEW WITH MR. GHEE LENG OOI, FOUNDER OF AKIRAKAN



**Ghee Leng, Ooi**  
**Co-Founder, AkiraKan**

## **Company Profile:**

AkiraKan is a dynamic and innovative company specializing in geospatial intelligence, intelligent transportation systems, and weather technologies. With a strong focus on leveraging edge compute, advanced algorithms and AI technology, they provide cutting-edge products and services that empower organizations across various industries. Via human-AI workflow integration, AkiraKan enables clients to make informed decisions, optimize operations, and gain valuable insights. Their collaborative and inclusive work culture promotes open communication and a sense of ownership among team members, ensuring a holistic approach to problem-solving. Committed to customer-centricity, AkiraKan delivers tailor-made solutions that address the unique needs and challenges of their clients. With a growing client base and

an expanding presence in global markets, AkiraKan is an up-and-coming player in the geospatial intelligence industry, driving innovation and shaping the future of spatial analytics.

## **What is AkiraKan all about and what are your main focus areas?**

AkiraKan is a company that was established in 2017, although our core team has been working together since 2012. Our primary goal is to provide continuous value creation for our clients. In 2012, we recognized the growing importance of artificial intelligence, but also identified significant obstacles in its implementation. People faced various challenges throughout the entire process, from data collection and pre-processing to post-processing and refining the data for effective AI model integration. Additionally, ownership of this process was a concern for enterprises,

governments, and individuals alike.

At AkiraKan, we specialize in working with data to address these challenges. When clients discuss their needs, they often refer to how they want to implement their workflows. For instance, in the surveying industry, clients may want to verify specific field points and ensure they are accurately represented in the data. They also seek ways to improve data quality control and integrate their knowledge and understanding into their workflow. Our aim is to continually create value for our clients in each of these aspects, forming a comprehensive and holistic value stream. We are dedicated to constantly improving and innovating this entire process.

Currently, our focus revolves around geospatial intelligence, intelligent transportation systems, and weather technologies. We strive to support our clients in these fields by leveraging our expertise and offering tailored solutions that align with their unique requirements. Our mission is to empower our clients with advanced technologies, enabling them to optimize their operations and make informed decisions in an ever-evolving end user needs and technology landscape.

**You mentioned that your core team was formed back in 2012. Did you all come together during your tertiary education? What was the inspiration behind the conception of AkiraKan?**

Yes, our core team consists of individuals who met during our time at the Hong Kong University of Science and Technology. We had the opportunity to collaborate in the RoboCon, and it became evident that we shared a passion for edge computing and artificial intelligence. As trend enthusiasts, we recognized the immense potential of these two fields and their future impact. The trend we observed was that computing power was becoming more affordable, compact, and increasingly powerful. This development opened up new possibilities for edge computing, where many tasks could be

performed locally. Additionally, as artificial intelligence models continued to advance, we anticipated that one day we will be able to witness many amazing applications. Today, the AI era has finally started to mature.

Driven by our belief in the potential of edge computing and artificial intelligence, we came together as a team to establish AkiraKan. Our vision was to harness the power of these technologies and leverage their capabilities to create innovative solutions for our clients. We understood the transformative nature of these fields and sought to be at the forefront of their application, providing cutting-edge AI-based services and driving advancements in the workflow processes of our clients.

**AkiraKan has a strong emphasis on privacy and responsibility in the field of AI. Could you explain why these are key priorities for the company? Additionally, given that AI has been utilized in questionable ways, how does your company approach this issue?**

Indeed, you are correct in highlighting the significance of privacy and responsibility, which have been focal points for us since the inception of our company. We recognized a growing sense of loss of control in today's interconnected world, particularly with regards to the technologies offered by major companies such as Google and Meta. Shoshana Zuboff's book, published in 2019, succinctly captures the essence of the issue, termed Surveillance Capitalism<sup>1</sup>. It refers to the collection of user data through ostensibly free services like searches and social networking platforms, where users are not merely users in the traditional sense, but rather the raw materials. The true beneficiaries are the business advertisers and the platform providers who leverage this data to target a broader audience. This business model raises ethical concerns and diminishes individual control over personal data, resulting in a pervasive feeling of powerlessness.

Our vision stems from this recognition, focusing on privacy and responsibility as fundamental principles. Our goal is to restore data

ownership to individuals and return to a more equitable business model. We envision a scenario similar to purchasing a car: if the car malfunctions, the company addresses and resolves the issue. However, with current products and services, users lack the ability to hold companies accountable or request changes. Similarly, when inaccurate information or fake news surfaces through search results, users are left with minimal control. Our aim is to reclaim this control.

To achieve this, we advocate for the implementation of an entire system that allows the users to have greater influence and decision-making power over algorithms and AI. Rather than relying on off-the-shelf AI solutions on cloud, which may yield uncertain outcomes despite rigorous training, we strive to provide customized and transparent AI models. By owning the entire process and enabling clients to shape the workflow, our services assist in fine-tuning algorithms, embedding cultural ethics and understanding, and empowering clients to take ownership of the process. This approach enables the enterprises and government entities we collaborate with to be accountable to their citizens and themselves, prioritizing responsible AI practices over potential data breaches. This commitment forms the foundation of our work.

### **So when you are providing your solution or product to the company, would your approach also include educating them about the ins and outs of it?**

When providing our solutions or products to companies, our approach goes beyond mere education. We engage in multi-lateral communication, taking into account both the management level and the working level within the organization. At the management level, we focus on discussing ethics, objectives, and the underlying philosophy related to their desired deployment. On the working level, our communication revolves around the practical aspects of the workflow, determining how to progress from point A to point Z. It is crucial for us to continuously enhance the integration of

management's ethics, philosophy, and objectives into the workflow. This collaborative approach ensures that our solutions align closely with the company's needs.

[So it is not just purely giving a product?]

Exactly. In today's context, providing a generic product without customization is highly unlikely to meet the real value and expectations of our clients. True value lies in tailoring solutions to match a company's unique work culture and requirements. This involves considering all aspects of their operations, rather than simply presenting them with a one-size-fits-all product. We refer to these tailored solutions as modules, which can be further developed and expanded over time. For example, our foundational AI models are pre-trained using mostly open data. From these models, we can deploy customized versions for different entities, ensuring that their private data remains exclusive to them, along with the AI models operating on their data. Our service encompasses the entire workflow, starting from understanding of the data and sometimes even data collection, e.g., real-time monitoring of traffic intersections. We build edge systems to assist in data collection, design APIs for seamless integration with back-end systems, and provide solutions for resilience and fallback servers if needed. Some clients may require on-premises hosting, and we help set up and maintain those servers as well. Ultimately, our approach revolves around catering to the specific needs of each client.

[Basically, you design a very multi-adjustable solution to however the clients would want it? So it's quite flexible in handling whatever their problem is.]

Precisely. We aim to provide a highly adaptable and flexible solution that can be adjusted according to the unique requirements of each client. The key elements of our approach involve the utilization of foundational AI models and provision of seamless workflow integration between human and AI. This allows us to create a truly customized solution that addresses the unique challenges and goals of our clients.

### **What is the current flagship product or solution of the**

## company to head towards the vision?

Our current flagship product that aligns with our vision is called "GPT-in-a-box." With the recent emergence of technologies like ChatGPT, it has become easier to understand the concept. We have been working on this type of solution for quite some time now. GPT-in-a-box represents a product that resides on the edge, owned by the client, and accompanied by foundational AI models that have a more generic nature. However, we assist our clients in tuning these models to suit their specific needs.

For example, let's consider a scenario where a client possesses roadside photo of a city. While the entire city may not have been scanned in detail, the client may be interested in monitoring the growth of trees and ongoing construction activities along the roads. These clients are interested in tracking changes within their environments. Additionally, they may already have a significant amount of related data stored in various files and folders. Our flagship product, GPT-in-a-box, can be seamlessly deployed and integrated into their existing system. Using the product, clients can simply input queries like "I want to know about the trees on this particular street" and generate an SQL query that retrieves the relevant data from the system. However, to ensure clarity and accuracy, we also offer additional services, such as collecting more data to train and improve the model.

In summary, our flagship product, GPT-in-a-box, encapsulates the core value we provide to our clients. It empowers them to leverage AI capabilities while retaining control over their data and tailoring the solution to their unique requirements.

**Considering that your product or service has many different technologies involved depending on the service, then your team would have to comprise of different areas of expertise?**

Indeed, the composition of our team consists of individuals with diverse areas of expertise. None of the core team members come from a Computer Science background. Instead, we have backgrounds in fields such as Civil Engineering, Electronics Engineering, and Industrial Engineering. While this may not have seemed apparent a decade ago, we now recognize that this is the direction things are heading. Previously, job roles and specific areas of expertise were more siloed. However, with the advancement of AI models and the integration of different technologies, the notion of rigid job roles is becoming obsolete. As we combine various AI models, the work becomes more interdisciplinary.

For example, the GPT models we deploy for clients can handle backend queries and even provide suggestions for database design. This merging of tasks and responsibilities means that individuals cannot simply focus on a single specific skill set anymore. Instead, we train and hire people who understand the importance of being versatile and adaptable in this rapidly evolving landscape.

While I hesitate to use the term "expertise," everyone on the team brings their own unique talents and abilities. Some individuals may excel in front-end development due to their flexibility and willingness to embrace change. They understand that they shouldn't be attached to a single role and are comfortable engaging in rapid prototyping and R&D. Interestingly, these individuals also work with AI and backend systems, which highlights the evolving nature of our field.

Overall, we prioritize working with individuals who are proactive in staying ahead of the latest trends and developments. We encourage them to explore new technologies and constantly challenge themselves. Rather than placing excessive emphasis on specific talents within a particular field, our focus lies in fostering personal growth and continuously expanding skill sets.

**The focus seems to be that when you're building a team is a bit more on their personality and character. That they have flexibility, and they would look for what's the latest thing and eager to learn instead of saying, I have how many years of**



## experience in this specific technology?

Yes, you're absolutely right. When building our team, we place a greater emphasis on an individual's personality and character rather than solely focusing on their years of experience or specific technologies they have worked with. Nowadays, experience in a particular technology can become obsolete quickly, so it's essential to have a mindset that is open to continuous learning and staying up-to-date with the latest advancements.

One of the key qualities we look for in potential team members is their ability to reflect on themselves and grasp the essence of things. This is precisely what the name AkiraKan represents. AkiraKan is a combination of two words pronounced in Japanese. The first word is "Akira" (諳), which translates to "essence," and the second word is "Kan" (観), meaning "to watch" or "to observe." Our intention is to pierce through the fog of superficial details and truly understand the core essence of our work. By doing so, we can continuously improve and make meaningful progress. This aspect of introspection and observing the essence is deeply rooted in our humanity.

It's not merely about calculations or technical skills but also about the ability to quiet down, reflect, and meditate on the right things and the true essence of technologies. While it's easy to get caught up in the rush of constantly moving forward, we value individuals who can find balance and engage in thoughtful contemplation. These are the qualities we seek in our team members, as they contribute to a culture of continuous improvement and a deep understanding of the work we do.

## Q8: It's a very interesting outlook. So with that, how do you identify people for your team since it's quite abstract and not just based on a CV?

Indeed, our approach to identifying people for our team goes beyond just evaluating their

CV or their current skill set. The interview process plays a crucial role in this regard. During the interview, we focus less on assessing their knowledge of current technologies or their proficiency in specific tools, and more on understanding their ideas and perspectives about the future.

We delve into discussions about their thoughts on the shortcomings and limitations of existing approaches and technologies. This allows us to gauge their ability to reflect on the current state of affairs and identify areas for improvement. We seek individuals who can think critically and envision innovative solutions that push the boundaries of what is currently known or practiced.

Additionally, while the interview provides valuable insights, it's also essential to work with the candidates in practical scenarios to truly understand how they can fit into our team. This could involve collaboration on projects or assignments that test their problem-solving skills, teamwork, and adaptability.

By adopting this approach, we aim to identify individuals who not only possess the technical skills but also demonstrate a deep sense of reflection, an ability to challenge existing norms, and a capacity to contribute to our team's growth and development.

## Part of your work culture has to do with a relatively close-knit group of people instead of departmental work where everyone does their own thing and in silos?

Absolutely, we strongly believe in fostering a close-knit work culture rather than promoting silos and departmental segregation. Over the years, we have realized the pitfalls associated with siloed work environments, and we strive to create a collaborative atmosphere within our team.

In a siloed setting, individuals tend to become defensive and protective of their own work, leading to a lack of openness and constructive feedback. However, in our approach, we encourage a culture where everyone is aware of the assumptions, limitations, and progress of the work being done. This shared understanding allows for

open discussions and constructive criticism.

By cultivating a culture of collective ownership, we create an environment where individuals feel comfortable expressing their concerns or suggestions. There is no sense of possessiveness or exclusivity over specific modules or areas of work. Instead, we value collaboration and collective problem-solving.

In our close-knit work culture, team members are encouraged to seek clarification and engage in discussions with others who are working on different aspects of the project. This promotes cross-functional learning and knowledge sharing. Everyone is empowered to contribute their ideas and perspectives to enhance the overall outcome.

With the rapid pace of technological advancements, it is crucial to have a culture that embraces flexibility, adaptability, and an open mindset. By breaking down silos and fostering a close-knit work environment, we can effectively navigate the challenges and capitalize on the opportunities that arise in our dynamic industry.

### **Is it difficult to maintain such culture as we commonly meet people who want to take the ownership as they put value in the things they do? Is it difficult to find likeminded people with this culture that you have?**

Maintaining our culture of collective ownership and collaboration does come with its challenges, especially when it comes to individuals who have a strong inclination towards taking ownership and attaching value to their work. However, we have found that the nature of the work we do and the unique challenges we face help foster the desired culture.

The work we undertake is often groundbreaking and novel, which means there are limited resources or pre-existing solutions available. We can't simply rely on copying and pasting or reusing modules. The scale of the data we work with is vast and often unstructured, requiring a deep understanding of the data to achieve optimal results. This

complexity extends beyond just AI models to encompass the entire workflow, including the user interface that reflects the data output. Understanding the interconnectedness of all these components is crucial.

In this context, each team member quickly realizes the humbling nature of the work. There are no easy shortcuts or plug-and-play solutions. It requires continuous communication, collaboration, and a willingness to learn and adapt. As individuals work on projects, they begin to grasp the importance of breaking down silos and engaging with others. They understand that assumptions can lead to pitfalls, and by actively seeking input and diverse perspectives, they can avoid potential issues and improve the overall outcome.

While it may initially require some adjustment, we find that individuals naturally transition into our open mindset as they engage in the work and witness the benefits of collaboration and collective problem-solving. The realization that everyone's contributions are valuable and necessary for success takes hold.

In summary, while there may be initial challenges in finding individuals who align with our culture, the nature of the work and the inherent complexity involved gradually shift people's mindsets. The humbling nature of the tasks at hand and the need for comprehensive understanding and collaboration naturally cultivate the desired culture of collective ownership and open-mindedness.

### **Would you consider this as a similar type of culture to research culture? Like when we are doing research work, we would want to share it and get comments, reviews and opinions from different parties and continuously work on the research. Would this be similar?**

Yes, there are similarities between our culture and research culture, but we approach it from a different perspective. In research culture, there is a top-down approach where managers and leaders hold the authority and make decisions about the

direction of the research. However, in our culture, we adopt a bottom-up approach.

Instead of having managers dictate the course of action, we empower the individuals executing and implementing the technical work to have the most influence. They become the drivers of decision-making and have the autonomy to shape the direction of projects. These individuals are the ones who possess the first-hand knowledge and insights about the work being done. They have a deep understanding of the client's needs, the architecture of the company, and the available resources.

By flipping the traditional hierarchy, we create a system where those closest to the work have the most say. They can advocate for the necessary resources, negotiate with clients, and provide insights into the shortcomings, possibilities, and potential of the product. This approach ensures that decisions are informed by the expertise and experience of those directly involved in the technical implementation.

We believe that this bottom-up approach avoids the obstacles and inefficiencies that can arise when decisions are made solely by managers who may be detached from the practical aspects of the work. It allows for a more agile and responsive environment where the collective wisdom of the team drives the progress and success of projects.

In summary, while there are similarities to research culture, our approach distinguishes itself by empowering the individuals executing the work to have a greater influence and decision-making power. This bottom-up approach leverages their expertise and enables a more efficient and effective development process.

**So it is a rather different type of approach in many different areas, even in companies. Would you consider this the point that is giving your company an edge in the industry?**

Yes, I believe our unique approach gives us

a distinct edge in the industry. The absence of a gap between our team members and the ability to communicate quickly and effectively with those at the working level is a significant advantage. Our bottom-up culture allows us to understand and respect the expertise of our team members who are directly involved in the technical work. They provide crucial information to the team in a timely manner, enabling us to offer concrete solutions to our clients.

When clients engage with us, they quickly realize that we can understand their needs and deliver tangible results. Our approach goes beyond simply pushing a product or deploying an API; instead, we establish a close connection and bridge any perceived distance. This level of engagement and understanding enhances our client relationships and builds trust.

Moreover, our culture fosters agility in our products. While the essence of a product remains relatively stable, we have the flexibility to adapt and fine-tune its application based on evolving requirements. This internal mindset encourages continuous reflection and ensures that we focus on the most relevant and impactful aspects of our work. As a result, our team members find their work more engaging and enjoyable, and we consistently strive to deliver the "right" solutions.

In summary, our distinctive approach not only enables effective communication within our team but also strengthens our client relationships and enhances the agility of our products. This combination of factors sets us apart in the industry, giving us a competitive edge.

**Would that be actually an issue, because your culture is very close-knit culture, focusing on the people. You did mention earlier that you are working on expansion already of your company. So would that be a problem? Because you would want to keep something close-knit but when more people come in, it tends to be a bit more difficult to maintain.**

We are aware of this potential challenge and have devised strategies to address it. While

maintaining a close-knit culture is important to us, we understand the need for expansion as our company grows. We have a plan in place to integrate new team members while preserving our culture.

We approach the integration process in a way that encourages collaboration and discourages silos. Those who have a deep understanding of our culture and values become torchbearers, guiding the new members. We organize ourselves into small units where everyone has access to information and resources, ensuring a seamless flow of knowledge and expertise.

When onboarding new team members, we assess their tendencies and talents and align them with specific areas. However, we also value cross-functional experience and periodically rotate individuals to different areas. This allows them to gain a holistic understanding of our technology and the bigger picture.

By striking a balance between maintaining our close-knit culture and integrating new members effectively, we aim to preserve our unique working environment while enabling growth and expansion. Our goal is to ensure that new team members can easily access the knowledge and resources they need, contributing to the collective success of the company.

### **For you as the co-founder of the company, and then you need communicate with everyone, wouldn't that take a lot of your time considering you need to be managing the whole company and so on.**

As the co-founder of the company, I believe that the time I spend communicating with everyone is a valuable investment. Each individual we hire is crucial and holds significance within our organization. We don't believe in creating tiers or hierarchies that limit autonomy and creativity. Instead, we provide a platform for each person to have autonomy and ownership over their work.

In our dynamic and ever-changing

industry, allowing individuals to have the freedom to create and innovate is essential. By fostering open communication with every team member, we cultivate a culture that values their ideas and contributions. These interactions are not a waste of time but rather a strategic investment in the core values and growth of the company and its individuals.

### **How do you see AkiraKan moving forward as a company? What is next?**

Moving forward, we see AkiraKan experiencing significant growth with the introduction of our GPT-in-the-box solution. While our focus remains on the geospatial domain, we are expanding into more generic applications. One of our key verticals is the intelligent transportation system, where we leverage geospatial data. Additionally, we are actively involved in the clean energy industry, combining geospatial work with renewable energy initiatives.

Another area of expansion for us is in weather forecasting and text analysis, where there is a high demand for accurate predictions. We aim to provide valuable insights by utilizing mapping technologies in various fields such as city management, rural planning, architecture, construction management, virtual asset management, and safety of critical infrastructures. These efforts mark the beginning of an exciting new chapter for AkiraKan, as we continue to evolve and broaden our impact in these verticals.

### **So, would it be like only Malaysia, Hong Kong or other places?**

Currently, we have clients in Germany, Singapore, Taiwan, and Switzerland. While we have established a strong presence in Malaysia and Hong Kong, our reach is not limited to these regions. As we continue to develop and expand our offerings, we anticipate attracting clients from a wider range of locations.

[So will foresee quite a lot of things from AkiraKan near future]

Our reputation and unique approach have led to organic outreach from interested parties.

However, we primarily operate as a B2B or B2G company, focusing on serving businesses and government entities. Our goal is to fulfil our clients' responsibilities to the people they serve by providing them with the necessary tools and ownership over their data. Therefore, while we may not engage in extensive marketing efforts, we remain dedicated to delivering value and meeting the needs of our clients.

**Please share some words of advice for individuals interested to join the industry, or perhaps even AkiraKan in the near future.**

When considering a career in the industry or potentially joining AkiraKan, there are a few words of advice I would like to share. First and foremost, bring a sense of fun and enjoyment with you. Embrace the journey and find joy in exploring new things. Second, maintain a level of dedication, not necessarily to a specific task or role, but to the excitement of working on innovative projects. Be open to new ideas and possibilities. Additionally, it is beneficial to have a general idea of what you want to achieve. While you may not have a clear direction initially, through active engagement, you may discover new perspectives and ways to improve existing processes. Finally, active participation is highly valued. Contribute your thoughts, ideas, and insights to the team. Your input can make a significant impact on the work we do.

---

# DATA AUGMENTATION STRATEGY FOR GENERATING REALISTIC SAMPLES ON DEFECT SEGMENTATION TASK

Massimo Martini, Università Politecnica delle Marche, Ancona, Italy

Riccardo Rosati, Università Politecnica delle Marche, Ancona, Italy

Luca Romeo, University of Macerata, Macerata, Italy

## Introduction

Machine Learning (ML) and Deep Learning (DL) are emerging as techniques for solving challenging tasks in challenging scenarios such as Industry 4.0 and precision farming.

Despite the advantages of these data-driven techniques, there are several disadvantages that could be overcome to scale these learning algorithms in a real-world scenario. One motivation is the difficulty of implementing an ad hoc data collection procedure to collect a supervised dataset that is highly representative of the production task to be solved. This strategy usually implies, for example, that in anomaly and defect detection tasks the number of good samples far exceeds the number of defective samples. As a result, there is a high imbalance between different classes in the classification or defect detection task. Downsampling strategies in this context aim to discard nondefect samples to maintain the balance between classes, thus drastically reducing the overall dataset. Several solutions to the class equilibrium problem have been proposed in the past, both at the data level and at the algorithm level. Data-level solutions include sampling methods such as different forms of resampling, e.g., random resampling, direct resampling (e.g., SMOTE), oversampling with informed generation (e.g., Generative Adversarial Network [GAN]). On the other hand, the algorithmic-level solution includes adjusting the costs of various classes (cost-sensitive learning), rebalancing the loss during the learning process, or automatically adjusting the threshold of the

posterior classifier. The sampling method could be easily scaled in the computer vision scenario as a data augmentation method to perform global augmentation (augmentation for the whole image) or to perform local augmentation (augmentation only for the region of interest).

In this paper, a data-level solution is introduced to improve the generalization performance of semantic segmentation of surface defects. This scenario can be relevant in different domains including the precision farming and agriculture scenario. Specifically, the proposed approach includes a generative phase to simulate synthetic defects and a validation phase to verify that the synthetic image is as close as possible to the real one. With real experiments on a benchmark dataset, we demonstrated the effectiveness of our approach in a real-use scenario compared to other widely used data augmentation approaches for semantic segmentation for defect detection. The data augmentation approach allows minority classes to be balanced while improving overall generalization performance.

## Methodology

The proposed approach consists of the following steps: preprocessing, data augmentation, defect validation and defect segmentation.

## Preprocessing

The first preprocessing step consists of

normalizing the pixels of the original images to values between 0 and 1.

Next, the images are divided into square patches of 512x512 pixels. The rationale behind image splitting and patch processing is mainly due to the main objective of the defect detection task, which is to maintain high image resolution, leading to very accurate and fine-grained segmentation.

Finally, a filtering step was added to retain only patches containing a distribution of defects. Since most of the image area mainly contains the background class, this step allows for a much more balanced dataset and to focus only on the candidate area where defects are present.

## Data augmentation

Two different approaches to perform data augmentation were considered: the first generates Gaussian defects completely independent of the dataset, while the second generates defects from the information obtained from defects already in the dataset.

Gaussian defects are applied randomly within the images of the dataset. They are configured according to several parameters: height and width of the two perpendicular Gaussian features, rotation angle, magnitude of the effect, type of defects

In contrast, the defects in the second approach are much more realistic because they take advantage of the defect information already present in the dataset. Given a generic defect and its mask, the entire portion of the defect is extrapolated. Then, in relation to the surrounding background, the defect is subtracted from it to obtain only the magnitude of the defect on the image. This "magnitude" can then be applied anywhere on another image. Defect generation was done using an algorithm that allowed a generated defect to be inserted within the images of the dataset, but without overlapping with existing ones.

For the generation of defects in the second approach, the following steps were followed: a defect is randomly chosen from existing ones; then the defect is extracted from its background, also based on its type; then, it is inserted into a random location of a random image, without

obtaining overlaps with those already present.

Before insertion, an additional processing step is carried out to obtain an increasingly realistic defect.

To the extracted defect, the following will be added:

- a Gaussian distributed additive noise;
- a Poisson-distributed noise generated from the data;
- a multiplicative noise
- no noise.

Since the former data augmentation approach allows the number of defects to be significantly increased, the latter ensures that all generated defects are as similar as possible to the original ones. Consequently, the proposed approach of validating the generated defects allows all unrealistic defects to be filtered out.

## Defect validation

If the generated defects do not reflect the distribution of the real ones, instead of improving segmentation, they may lead to the opposite result. For this reason, it is necessary to use a method that can validate the performed data augmentation procedure and thus the generated defects, so as to discard low-quality defects that deviate from the original ones.

The approach proposed in this paper is based on the concept of Siamese Networks, which allows learning a similarity function that, given two inputs, can evaluate how similar they are, as a function of a numerical value.

Siamese networks are neural networks that share weights between two or more sister networks, each of which produces latent vectors of their respective inputs.

In supervised similarity learning, networks are trained to maximize the contrast (distance) between latent vectors of inputs of different classes and minimize the distance between embeddings of similar classes, resulting in latent spaces that reflect the class segmentation of the training inputs.

The Siamese network in this context is used to understand the quality of the generated defects, so that only the "best" ones (closest to the real ones) are filtered out and then used for training the segmentation network.

## Defect segmentation

The network for semantic defect segmentation

used in this work is based on the well-known UNet network. The network was pre-trained on the Imagenet dataset and then re-trained with the dataset proposed in this paper. The loss function used for training the network is a combination of two very common losses in the semantic segmentation task (Dice and Focal).

## Results

For the evaluation of the proposed method, a new dataset was collected using an RGB camera. A single view of several plastic objects was acquired with the main purpose of detecting structural defects on rubber materials.

Table 1 describes the results of the defect segmentation model in terms of IoU Score and F1-Score using the original dataset (Original), the dataset augmented with the proposed data augmentation (DA), and the dataset augmented and validated according to the proposed methodology (DA+Validation), respectively.

|                 | IoU score | F1 score |
|-----------------|-----------|----------|
| Baseline        | 82.11     | 83.97    |
| DA              | 82.12     | 84.19    |
| DA + Validation | 84.10     | 85.61    |

*Tabella 1: Results of the proposed DA methodology*

The extracted results highlight the potential of the proposed methodology in the specific use case (segmentation of plastic material defects) and the possibility of generalization also in a different scenario such as precision farming and agriculture.

## Reference

Jain, Saksham, et al. "Synthetic data augmentation for surface defect detection and classification using deep learning." *Journal of Intelligent Manufacturing* (2022): 1-14.