Received: Added at production | Revised: Added at production | Accepted: Added at production

DOI: xxx/xxxx

RESEARCH ARTICLE

Morphology-guided convolutional Graph Neural Network decodes optically barcoded nanoparticles for one-pot, purification-, amplification-, and enzyme-free femtomolar nucleic-acid diagnostics

Kartik Jhawar[†1,2] | **Xiao-Liu Chu**[†1] | **Joseph B. DeGrandchamp**[†1] | **Yueming Yin**[1,3] | **Lalitha Anjali Peddibhotla**[1,4] | **Shabina Banu**[1,4] | **Farah Najwa Nabila Mohd Hatta**[1,5] | **Yaoli Wang**[9] | **Peter Török***[1,8,6,7] | **Lipo Wang***[1,2]

[1]Institute for Digital Molecular Analytics and Science, Nanyang Technological University, 636921, Singapore

[2]School of Electrical and Electronic Engineering, Nanyang Technological University, 639798, Singapore

[3]College of Computing and Data Science, Nanyang Technological University, 639798, Singapore

[4]Republic Polytechnic, 738964, Singapore

[5]Universiti Putra Malaysia, 43400, Selangor Darul Ehsan,Malaysia

[6]Lee Kong Chian School of Medicine, Nanyang Technological University, 636921, Singapore

[7]School of Physical and Mathematical Sciences, Nanyang Technological University, 637371, Singapore

[8]Singapore Centre for Environmental Life Sciences Engineering, Nanyang Technological University, 637551, Singapore

[9]College of Electronic and Information Engineering, Taiyuan University of Technology, Taiyuan, China

Correspondence

† Authors contributed equally to the article.

Corresponding author(s) Peter Török and Lipo Wang

Email: peter.torok@ntu.edu.sg, elpwang@ntu.edu.sg

Abstract

Affordable, accurate, and rapid point-of-care diagnostics remain elusive because existing nucleic acid tests generally require amplification, enzymes, purification, or prohibitively expensive equipment. Here, we present a one-pot assay that eliminates such requirements by pairing silver and gold nanoparticles with distinct light-scattering properties as optical barcodes, forcing hetero-particle pairing and allowing target-induced clustering to be imaged on a low-cost dark-field microscope, without costly fluorescent labels or bulky instrumentation. A novel deep learning architecture, Morphology-guided convolutional Graph Neural Network (Mc-GNN), combines morphology-guided kernel convolution and graph-based context-aware relational modelling to feature extract intuitively on single unbound particles and small cluster tensors, processing concurrently across ∼5000 particles per field using only <5 GB consumer GPUs. Mc-GNN, an ablation-validated novelty, near-perfectly classifies images by concentration down to femtomolar levels, with 98.2% average recall for the detection of synthetic DNA and performs robustly (94.8%) for RNA detection from whole SARS-CoV-2 virus, even with variations in nanoparticle selection and sample complexity, surpassing all baselines and popular architectures ($\leq$ 90% for viral RNA). This computationally efficient, smartphone-compatible approach is readily extensible to new analytes and multiplexing, offering a scalable solution for fieldable diagnostics and future applications.

KEYWORDS

Deep learning, Dark-field microscopy, Nanoparticle, Diagnostics, SARS-CoV-2

1 | INTRODUCTION

The start of the COVID-19 pandemic clearly highlighted the limitations of the available infectious disease diagnostic tools. Conventional nucleic acid amplification methods (e.g., RT-PCR) provided high sensitivity, but were hampered by high costs,

Abbreviations: NP, nanoparticle; AuNP, gold nanoparticle; AgNP, silver nanoparticle; Au@SiNP, gold-shelled silica-core nanoparticle.

lengthy protocols, and centralized laboratory requirements[1,2]. Antigen and serological assays, while deployable outside laboratory settings, often compromised sensitivity and specificity, particularly during the early infection stages when virulence is high[2]. Consequently, there was, and still is, an urgent clinical need for rapid, affordable, and reliable point-of-care (POC) diagnostics that require minimal infrastructure and training.

Lateral flow assays (LFAs), exemplified by widely used COVID-19 antigen tests, partially meet these criteria by offering simplicity and affordability. However, they suffer from intrinsic sensitivity limitations, leading to high false-negative rates for low-abundance targets in the early stages of the disease[3,4]. To improve sensitivity beyond LFAs, nanoparticle-based colourimetric assays have been explored extensively. These assays rely on analyte-driven aggregation of noble-metal nanoparticles (NPs), whose localized surface plasmon resonances (LSPRs) shift upon particle clustering, amplifying the detection signal. Early seminal work demonstrated DNA-induced gold NP aggregation can indicate particle cross-linking via distinct optical changes[5,6]. Yet conventional NP aggregation assays have fundamental limitations: detectable LSPR shifts diminish sharply once interparticle distances exceed $\sim$10 nm, making such assays effective only for targets that bring NPs into extremely close proximity. This distance-dependent sensitivity restricts conventional plasmonic readouts to particular molecular geometries and binding configurations.

Researchers have pursued various strategies to overcome these limitations, including employing multiple spectrally distinct NPs to generate readouts independent of sub-10 nm plasmonic coupling. Hetero-nanoparticle systems producing multi-colour optical signals upon target binding effectively sidestep the coupling limit by encoding detection in co-localized but spectrally separate NP signals[7–9]. While these optical encoding strategies have achieved impressive multiplexing capabilities, they often require complex multi-component designs (e.g., lipid-bilayer circuits or elaborate colour codebooks) and do not leverage direct image-based analysis of individual NP events. Alternative approaches integrate NP assays with image-based detection and machine learning, using smartphone cameras or simple microscopes coupled with classical algorithms for analyte estimation[10,11]. Single-particle dark-field imaging combined with CRISPR-based assays has also visualized NP clustering events directly[12]. However, these implementations typically depend on coarse bulk-colour features, manual counting, or simple classifiers, lacking particle-level morphological insight essential for high-sensitivity diagnostics[13–18]. Recent deep learning (DL) advancements using convolutional neural networks have further improved sensitivity but mostly operate at the whole-image level, requiring external labeling, predefined ROI alignment, or significant computational resources[10,11,19–25]. Only isolated cases like Shiaelis *et al.*[21] have analyzed individual NPs with DL, yet still necessitate fluorescent labeling and labor-intensive ROI alignment.

Here, we bridge the gaps by integrating (1) image-based single-NP detection, (2) the use of orthogonally coloured NP scatterers, and (3) a morphology-guided DL architecture for high-throughput, intuitive particle-level feature extraction. This three-pronged strategy dramatically enhances the sensitivity and reliability of NP aggregation diagnostics. Using an economical, low-resolution dark-field microscope, our assay directly captures analyte-dependent NP co-localization events without any fluorescent labels or bulky spectroscopic equipment, enabling compatibility with portable or smartphone-based microscopy setups[11,19]. Crucially, deploying orthogonally coloured NPs provides an optical "barcode" through enforced hetero-particle pairing upon target binding, acting as a built-in specificity filter against spontaneous aggregation. And in contrast to earlier hetero-particle implementations[7–9], our method simplifies the assay into a one-pot, purification-free protocol, retaining ease-of-use while laying the groundwork for full multiplexing in future iterations. Notably, unlike for standard nucleic acid tests such as RT-PCR[26,27], this one-pot design is feasible because the detection chemistry is entirely protein-free, allowing for incorporation of lysing reagents directly for the release nucleic acid targets without inhibiting performance[28–30].

Finally, at the core of this combination is our novel DL architecture, Mc-GNN, which powers our analytical framework. Essentially, Mc-GNN combines: (i) learnable convolutional kernels explicitly pre-multiplied by morphological masks (circular, Gaussian, donut, square, and edge), intuitively encoding particle geometry and scattering patterns at the single-particle level; (ii) subsequent graph modelling and graph convolutional layers that perform context-aware relational reasoning and feature aggregation using the morphology-guided features extracted for each particle, thereby enhancing feature reasoning; (iii) efficient computation on low-memory, consumer-grade GPUs (requiring less than 5 GB RAM and supporting parallel processing of up to 5,000 particles); and (iv) an integrated framework that robustly discriminates analyte concentrations down to femtomolar (fM) levels, taking into account all unbound particles and small clusters without the need for fluorescence labelling or region-of-interest (ROI) alignment [19,21]. This architecture markedly outperforms contemporary approaches in quantifying concentrations of synthetic DNA and SARS-CoV-2 viral RNA samples, demonstrating exceptional accuracy, computational efficiency, and assay robustness. Current morphology-guided convolutional kernels[31–33] primarily target computationally intensive semantic segmentation tasks[34–42], limiting particle-level interpretability. Graph-based approaches, such as KerGNN[43], dual-graph CNNs[44,45], and hybrid CNN-GNN frameworks (e.g., SR-GNN, CGC-Net, MorphoGNN[46–49]), as well as classic GNNs[50–52],

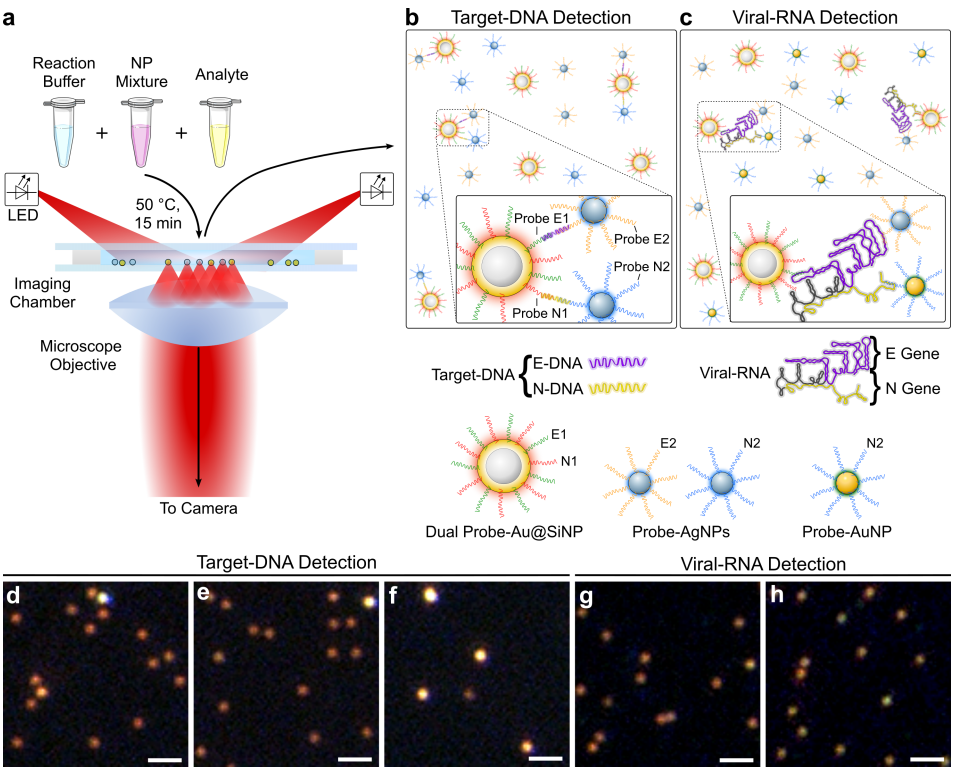

FIGURE 1 Nanoparticle imaging assay overview. **a** A schematic representation of the assay steps, with optically barcoded nanoparticles (NPs), analysed via dark-field microscopy. **b,c** Schematic representations of nanoprobes binding to Target-DNA (mixed ssDNA derived from SARS-CoV-2, see Table S3) or Viral-RNA (genomic RNA from whole SARS-CoV-2), respectively. **d-f** Exemplary dark-field images (RGB) of the Target-DNA detection scheme, showing images captured for control (**d**), 500 fM Target-DNA ([E-DNA] = [N-DNA] = 500 fM) (**e**), and 50 pM Target-DNA ([E-DNA] = [N-DNA] = 50 pM) (**f**). **g,h** Exemplary dark-field images (RGB) of the Viral-RNA detection scheme, showing images captured for control (**g**) and 445 fM Viral-RNA (**h**). Scale bars = 5 µm.

similarly lack explicit morphology-aware constraints at the initial convolution stage. In contrast, our proposed Mc-GNN uniquely integrates morphology-guided kernel convolution with relational graph modeling, enabling unified spectral (scattering/emission), spatial (distribution/spacing), and morphological (size/shape) nanoparticle-level learning to address critical gaps in NP-biosensing workflows.

By merging affordable optical imaging, NP spectral differentiation, and an intuitive DL architecture, our Mc-GNN-powered assay addresses longstanding limitations in POC diagnostics, providing a scalable, fieldable solution that bridges laboratory performance with real-world diagnostic needs.

2 | RESULTS AND DISCUSSION

In this section, we present a NP aggregation assay formulated for nucleic acid detection, using our proposed DL model to quantify the target molecule concentration (referred to as the "Concentration Prediction" task). First, we characterize the NP functionalization and image analyte-dependent NP cluster formation at a series of synthetic DNA concentrations via dark-field imaging. Next, we develop a DL model Mc-GNN to determine DNA concentration and then evaluate its classification performance. We compare our method's efficacy with existing approaches and other popular models, and finally, apply the strategy with a one-pot assay that directly probes SARS-CoV-2 viral material from heat-inactivated whole virus.

2.1 | Nucleic acid detection assay and raw data overview

In a diagnostic test based on imaging NP aggregation, the challenge is to correctly differentiate single unbound NPs from clusters formed in the presence of target molecules. The concentration can therefore be inferred by counting the different particle sub-populations. However, low target molecule concentrations result in a population that mainly consists of single unbound particles and small clusters. Due to the diffraction limit, these are difficult to differentiate optically based on the information derived from their size and scattering intensity alone, whereas spectral data provides an additional information channel for exploitation. As displayed in Figure 1, we employed gold-shell silica-core NPs (Au@SiNPs) as prominent red scatterers, spherical silver NPs (AgNPs) as blue scatterers, and, in our final assay iteration, spherical gold NPs (AuNPs) as additional green scatterers to best utilize a Red / Green / Blue (RGB) colour space (Figure S2). This enables co-localization of "optically barcoded" nanostructures within the same diffraction-limited point spread function (PSF), imaged with a colour (RGB) camera integrated in a custom dark-field microscope, with example images seen in Figure 1d-h. By illuminating the samples with broad-spectrum light-emitting-diodes (LEDs) as an affordable light source, we observe subtle but measurable variations in the RGB intensity profiles of NP clusters versus single particles. To generate labeled training data for the DL model, we implemented a nucleic acid detection assay by functionalizing our NPs with complementary DNA handles, and we systematically captured thousands of images of particles in an automated fashion using a piezo stage.

To inform our initial assay design, we first exploited a combined equimolar mixture of a synthesized 188-bp single-stranded DNA (ssDNA) ("E-DNA") derived from SARS-CoV-2 Envelope gene (*E* gene) and a 200-bp ssDNA ("N-DNA") from SARS-CoV-2 Nucleocapsid gene (*N* gene), that we will henceforth refer to as "Target-DNA" ([Target-DNA] = [E-DNA] = [N-DNA]). By mixing the two ssDNA sequences, we attempted to approximate genomic RNA detection (where both genes are present). Despite obvious limitations, this choice allowed us to test the assay performance with well-characterized and stable target molecules before attempting to detect full-length SARS-CoV-2 RNA. It also simplified probe selection, as the same DNA probes could be used for both applications. For Target-DNA detection, we used Au@SiNPs and AgNPs only, with probes complementary to E-DNA and N-DNA (Figure 1b). The NPs were first mixed with a buffer formulated to reduce non-specific aggregation before adding Target-DNA (or blank buffer) at defined concentrations spanning an order of magnitude (50 fM, 500 fM, 5 pM, and 50 pM). Samples were then heated briefly (15 min) at 50 °C before transfer to an imaging chamber for dark-field analysis. We utilize the beneficial characteristics of Au@SiNPs, which due to their larger mass will settle on the glass coverslip surface naturally by gravity. The Au@SiNP spots can then be easily localized and further analysed for their spectral characteristics to determine binding of AgNPs in a Target-DNA dependent manner (Figure 1d-f).

We then further tested the limits of the assay with heat-inactivated SARS-CoV-2 virus (capsid intact) as target (henceforth "Viral-RNA") (Figure 1b). This task increases the complexity, as the virus is provided in serum-fortified cell growth media (which could destabilize NPs) and also requires viral lysis. However through incorporating high concentrations of surfactants into the assay buffer, we were able to simultaneously utilize the protective effect of surfactants for NPs[28,29] as well as their lysing potential[30], resulting in a one-pot assay format. In an effort to increase sensitivity, we also replaced one species of AgNP probe with AuNPs in the assay in order to fully utilize the RGB colour space (Figure 1c). As such, released Viral-RNA could potentially cross-link all three NP species, and as our Au@SiNPs are dual probe, they can ostensibly bind to the fully intact target or to *E* or *N* gene-containing fragments. Through dilution of the virus sample, we proceeded with the assay at virus concentrations of 111 fM, 223 fM, and 445 fM (genome equivalent concentration), generating dark-field images for downstream analysis (Figure 1g,h).

2.2 | Deep Learning Analysis Overview

To robustly quantify biomolecular binding events by capturing the spectral and morphological cues in NP-based microscopy assays, we propose Mc-GNN. For the Concentration Prediction task, integer labels were used to encode each concentration level. In preparation for particle-level feature extraction, the raw image datasets were first filtered to remove blurry or poorly focused images, and then processed through the preprocessing sequence as illustrated in the experimental pipeline overview Figure 2. Raw dark-field microscopy images were pre-processed in multiple steps performing, global background correction, CLAHE-based contrast enhancement[53], and Gaussian blur, which supports for proper contour detection. This is followed by particle-wise background correction on the raw image, by estimating local background in an expanded region around each particle while masking out nearby particles or clusters, along with radial sorting of the detected contours (from the optical center outward) and cropping of individual particles. Radial sorting prioritises particles close to the optical axis, which are least affected by field-curvature defocus and vignetting that degrade image quality toward the periphery[54,55]. The sorting across raw

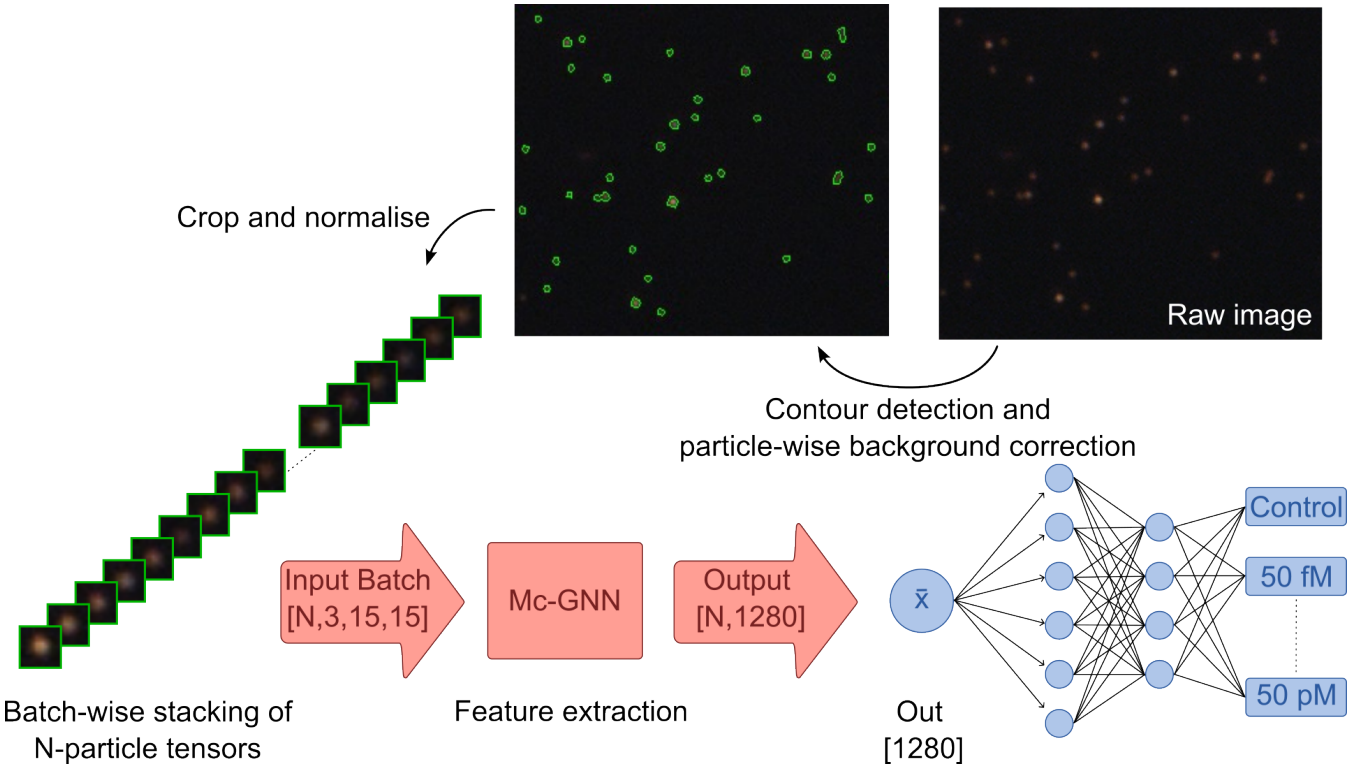

FIGURE 2 DL analysis experimental pipeline. Raw images undergo global correction to support contour detection (-steps not depicted in the Figure), followed by particle-wise background correction, radial sorting, and cropping. All the N-particle tensors are stacked in shape $[N, 3, 15, 15]$ and processed by Mc-GNN for feature extraction, followed by fully connected layers for the concentration prediction.

images also facilitates the incorporation of spatial information into the analysis. Particle-wise background correction further compensates for local illumination drift and detector offsets. Together, these steps provide uniformly focused, intensity-balanced particle crops and removes the need for an explicit ROI alignment stage. Each segmented particle is resized into a uniform tensor of shape $[3, 15, 15]$, and all particles from the raw image (typically the number of particles found, $N \approx 1,000\text{-}5,000$) are stacked in batched input tensors of shape $[N, 3, 15, 15]$.

These particle tensors are fed into the Mc-GNN model, architecture of which is shown in Figure 3. It is composed of two core components: (i) a morphology-guided convolutional stage that employs five types of structural masks-circular, Gaussian, donut, and square for morphology-guided feature extraction; and (ii) a graph-based relational modeling that models morphology-guided kernel outputs from step (i) as a Graph followed by the GCN (Graph Convolutional Network; a common layer in GNNs that aggregates neighbor features) layers. This results in relational learning of meaningful interactions between multiple morphology-guided features at the particle level through graph-based feature propagation. Thus, making it overall an intuitive model, as every structural mask highlights a distinct attribute that the model extracts from an individual particle and, following graph-based steps, does relational learning to enhance the extracted information. The resulting particle-level embeddings are aggregated into image-level representations and passed through a multilayer perceptron, often referred to as fully connected layers.

Model performance was evaluated using standard classification metrics, including Precision, macro Recall (same as average class-wise accuracy), and macro F1-score, averaged across 5-fold stratified cross-validation. For clarity, we omit the prefix "macro" in subsequent references to these metrics. AUC was excluded for the concentration prediction task due to its multiclass structure and class imbalance. To assess statistical significance and compare model performance, we first applied the Friedman test to the performance scores of all models across cross-validation folds, using averaged ranks (referred to simply as Rank; can be non-integer values as they are averaged across folds) to summarize relative performance. If the Friedman test yielded statistically significant differences (as indicated by the χ^2 statistic and p-value), we conducted Wilcoxon signed-rank post-hoc tests to identify which specific model pairs exhibited statistically significant differences[56]. We systematically evaluated the proposed Mc-GNN on both the prepared datasets. All models were trained using 5-fold stratified cross-validation with consistent

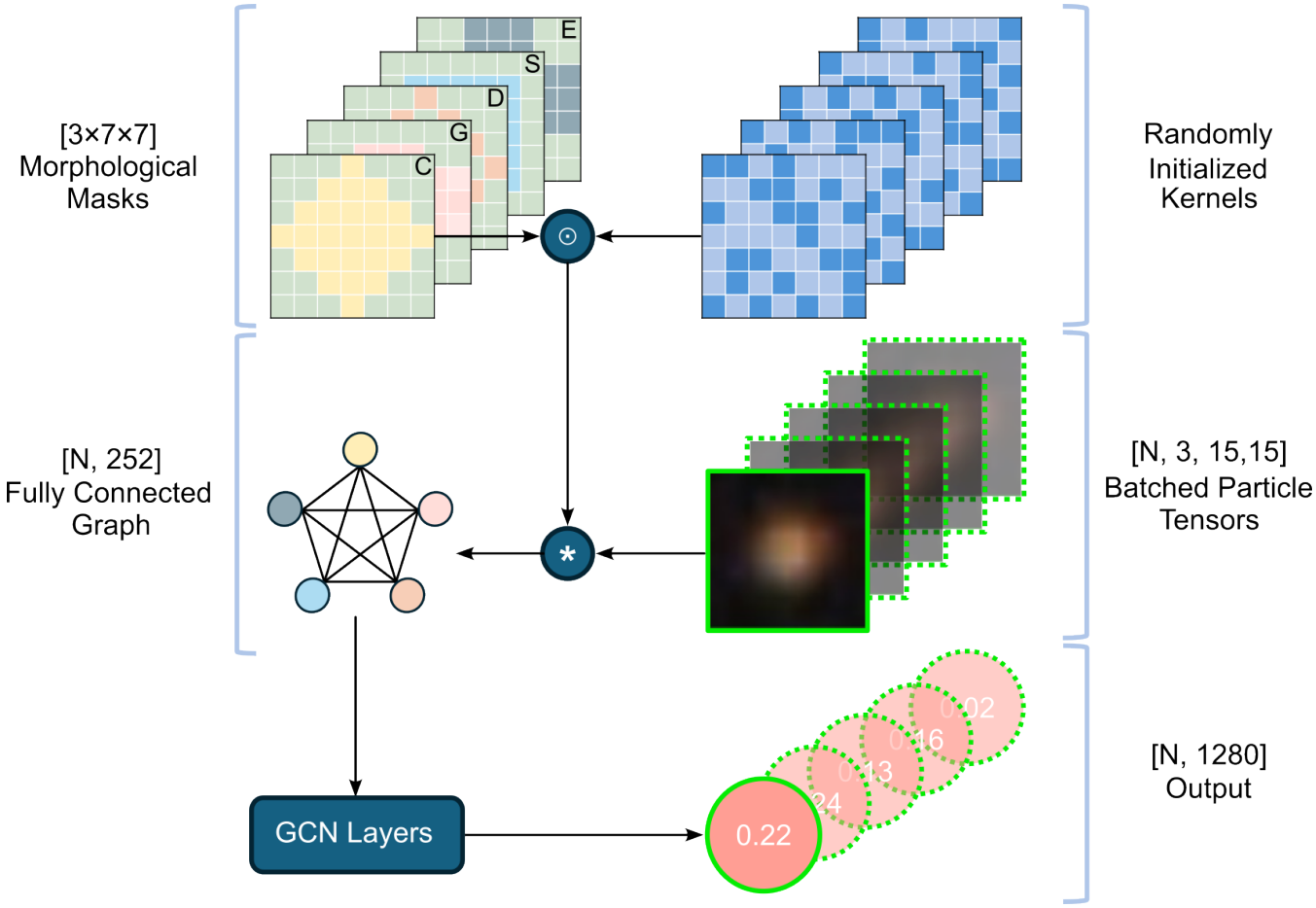

FIGURE 3 Architecture of the proposed Mc-GNN model. Each input raw image is preprocessed into batched particle tensors $[N, 3, 15, 15]$, which are convolved with randomly initialized kernels modulated by defined morphological masks - C: Circular, G: Gaussian, D: Donut, S: Square, E: Edge, 252 of each sized $[3 \times 7 \times 7]$. Spatially compressed resulting feature embeddings $[N \times 5, 252]$ form nodes in a fully connected graph, where graph-based relational learning across morphology-guided kernel outputs is performed via the GCN layers. This yields final particle-wise outputs $[N, 1280]$ for downstream predictions. "$\odot$" denotes Hadamard (element-wise) product; "$*$" denotes convolution.

preprocessing as described previously, data partitions, and hyperparameters. While we did not test additional random seeds, the consistent cross-validation results across folds suggested that the model's performance is stable and not highly sensitive to initialization as reflected by Friedman rankings of consistent 1.0 across ablation studies, and comparative performances [see Tables 1 2 3].

Figure 4 provides a comprehensive evaluation of proposed model performance across the Target-DNA and Viral-RNA datasets. In the best cross-validation fold (reported solely to illustrate the model's potential, while noting the class imbalance and the possibility of variation due to random initialization and inter-fold differences), the proposed Mc-GNN achieves high recall, with only 4 misclassifications in the Target-DNA dataset, accurately identifying 50 fM and 500 fM, while 98.42% control samples, 98.90% 5 pM samples, and 99.17% 50 pM samples (Figure 4a). In the Viral-RNA dataset (Figure 4b), the model correctly classifies 111 fM, while 95.88% control samples, 92.73% samples at 223 fM, and 91.72% samples at 445 fM. In addition to concentration prediction, we first evaluated all models on the clinically relevant analyte detection task, in which the control class (absence of target analyte=label 0) was compared against all non-control concentrations (analyte present=label 1). The corresponding best fold test results for this task are presented in Supporting Information (SI), Figure S1a,b.

We benchmarked Mc-GNN against a set of popular DL architectures, including ShuffleNet[57], MobileNet[58], GhostNet[59], EfficientNet-B0[60] (as also used by Jin *et al.*[24]; henceforth referred to as EfficientNet), and ResNet9[61](as higher variants (e.g., ResNet18, as also modified and used by recent McAffee *et al.*[25]) incurred prohibitive computational costs in our particle-level

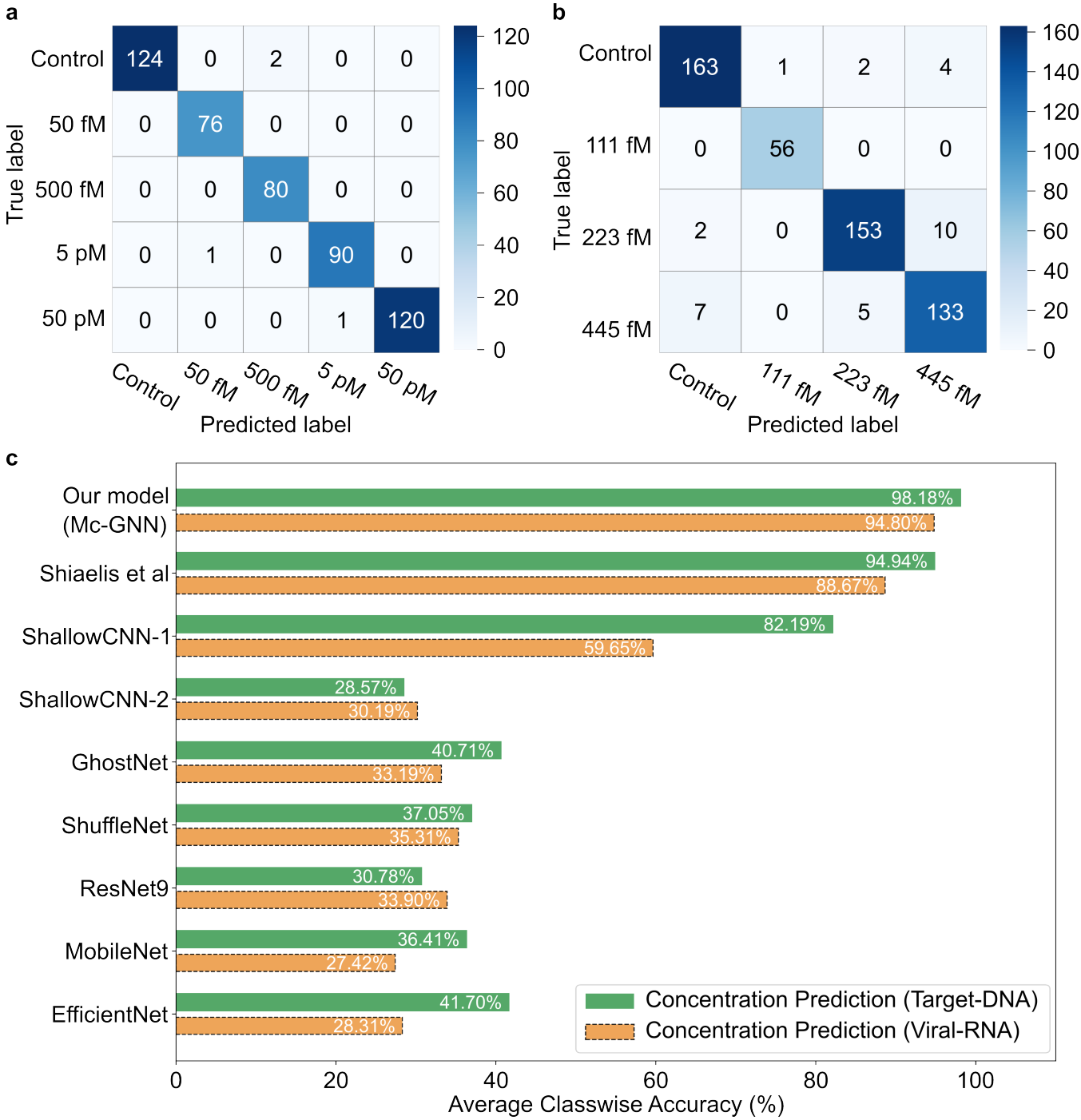

FIGURE 4 Best fold test results: **a** Confusion matrix for Target-DNA concentration prediction. **b** Confusion matrix for Viral-RNA concentration prediction. **c** Comparative performance of different models on both datasets.

setting), as well as domain-specific DL baselines relevant to this work. Among these, the model by Shiaelis *et al.*[21] is considered a strong contemporary baseline. The method proposed by Sajed *et al.*[11], which employs a simple architecture with a single Conv2D layer, is referred to as ShallowCNN-1 in our comparisons. Furthermore, we observed performance degradation when increasing convolutional depth beyond two layers; hence, we grouped the architectures from Liu *et al.*[22], Tun *et al.*[10], and Kang *et al.*[23] under a unified label *ShallowCNN-2*, as they follow a similar structural pattern involving two or more than two Conv2D layers. We did not benchmark against the AlexNet variant used by Wang *et al.*[20] or the adaptive FPN model proposed by Song *et*

al.[19], as both were designed for full-field microscopy images. Both the approaches were not compatible with our particle-level parallel processing framework and incurred prohibitive memory costs when adapted to our setting, which ultimately led us to exclude them from our comparisons.

Figure 4c presents a comparative accuracy analysis across multiple models. Mc-GNN consistently outperforms baselines in concentration prediction task, achieving 98.18% avg. classwise accuracy on the Target-DNA dataset, and 94.80% on the Viral-RNA dataset. In contrast, the next best-performing model, Shiaelis *et al.*[21], achieved 94.94% accuracy on the DNA dataset and 88.67% on the Viral-RNA dataset. Most other CNN-based baselines and popular models reported substantially lower performance, falling below 83% for Target-DNA and 60% for Viral-RNA, with some models even performing below 50%. These low accuracies suggest that these popular DL models may be inherently unsuitable for these datasets, rather than indicating implementation issues, as training curves converged and identical settings yielded substantially higher accuracies for the analyte detection task on both the datasets (see SI, Figure S1c). Given a brief exploration of alternative hyperparameters yielded no meaningful improvement to the baseline models for the analyte detection task, we refrained from extensive tuning here. Furthermore, unlike these baseline architectures, our proposed model intuitively incorporates explicit morphological guidance into its feature extraction process, which may contribute to more robust results. Notably, it also maintains computational efficiency, requiring less than 5 GB of GPU memory for training a batch size of 64 raw images, whereas most other DL architectures typically exceed 10 GB under similar training settings. Thus, the results highlight the efficiency and robustness of Mc-GNN across both modalities.

We also conducted ablation studies for the focused task of concentration prediction using the proposed Mc-GNN model, and the results are discussed ahead. The detailed model performance comparison analysis, along with the accompanying statistical analyses are presented in the subsequent subsection. This substantiates the consistent superiority of Mc-GNN for concentration prediction tasks. This initial diagnostic assessment provides foundational support for the subsequent concentration prediction results. Comprehensive results and comparative analyses for the analyte detection task are presented in the SI (Figure S1, Table S1,S2) for support.

2.3 | Ablation Study

To evaluate the contribution of individual architectural components in our proposed Mc-GNN model, we conducted an ablation study on the concentration prediction task using both the Target-DNA and Viral-RNA datasets. The baseline variants included: *(i)* Morphology-Guided Convolution, *(ii)* Fixed-Shape Kernel Convolution, and *(iii)* Fixed-Shape Kernels with GNN. The ablation study focused on the concentration prediction task, which forms the primary focus.

We evaluated all models across five stratified folds and reported the mean Precision, Recall, and F1-score along with Friedman rankings and Wilcoxon signed-rank test results against the full Mc-GNN model (see Table 1). Friedman tests confirmed statistically significant differences in model performance across all three metrics on both datasets: Target-DNA dataset: Precision ($\chi^2 = 13.56$, $p = 0.0036$), Recall ($\chi^2 = 15.00$, $p = 0.0018$), and F1-score ($\chi^2 = 14.04$, $p = 0.0029$); Viral-RNA dataset: Precision, Recall, and F1-score all yielded identical values ($\chi^2 = 14.04$, $p = 0.0029$). Across both datasets and all metrics, Mc-GNN achieved the best average rank (1.0), demonstrating the synergistic benefit of combining morphology-guided kernels and graph-neural network. Post-hoc Wilcoxon signed-rank tests revealed that Mc-GNN significantly ($p < 0.05$) outperformed all ablation variants across every metric, underscoring the importance of each design choice. Notably, obtaining the best rank in four out of five folds, but a lower rank or tie in the remaining fold, can yield a moderate yet statistically significant p-value (e.g., $p = 0.0312$). Such a scenario illustrates that even small but consistent improvements, accompanied by occasional ties or minor rank variations, can result in statistically meaningful outcomes in Wilcoxon signed-rank tests.

2.4 | Comparative Performance Analysis

As shown in Table 2, the Friedman test revealed statistically significant differences across the evaluated baselines on the Target-DNA concentration prediction task for all three metrics: precision ($\chi^2 = 31.36$, $p = 1.21 \times 10^{-4}$), recall ($\chi^2 = 32.96$, $p = 6.3 \times 10^{-5}$), and F1-score ($\chi^2 = 33.49$, $p = 5.0 \times 10^{-5}$). Our proposed Mc-GNN consistently achieved the top rank (1.0) across all metrics, delivering precision of 0.9813, recall of 0.9818, and F1-score of 0.9814. Post-hoc Wilcoxon signed-rank tests confirmed that Mc-GNN's improvements over each baseline were statistically significant ($p = 0.0312$ for all metrics).

TABLE 1 Ablation study of the proposed Mc-GNN on both datasets for the concentration-prediction task. Average 5-fold scores, Friedman rankings, and Wilcoxon signed-rank p-values (vs. Mc-GNN) are reported. Statistically significant differences ($p < 0.05$) are shown in **bold**.

Target-DNA Dataset

Model	Precision			Recall			F1-Score		
	Score	**Rank**	p-val	**Score**	**Rank**	p-val	**Score**	**Rank**	p-val
Mc-GNN	**0.9813**	**1.0**	–	**0.9818**	**1.0**	–	**0.9814**	**1.0**	–
Morphology-Guided Conv.	0.9636	2.0	**0.0312**	0.9663	2.0	**0.0312**	0.9643	2.0	**0.0312**
Fixed-Shape Kernels + GNN	0.9069	3.4	**0.0312**	0.9206	3.0	**0.0312**	0.9103	3.2	**0.0312**
Fixed-Shape Kernel Conv.	0.8805	3.6	**0.0312**	0.8799	4.0	**0.0312**	0.8729	3.8	**0.0312**

Viral-RNA Dataset

Model	Precision			Recall			F1-Score		
	Score	**Rank**	p-val	**Score**	**Rank**	p-val	**Score**	**Rank**	p-val
Mc-GNN	**0.9476**	**1.0**	–	**0.9480**	**1.0**	–	**0.9474**	**1.0**	–
Morphology-Guided Conv.	0.9302	2.2	**0.0312**	0.9266	2.2	**0.0312**	0.9279	2.2	**0.0312**
Fixed-Shape Kernels + GNN	0.8646	2.8	**0.0312**	0.8730	2.8	**0.0312**	0.8646	2.8	**0.0312**
Fixed-Shape Kernel Conv.	0.7039	4.0	**0.0312**	0.7212	4.0	**0.0312**	0.6953	4.0	**0.0312**

TABLE 2 Performance Comparison of Proposed Mc-GNN with Contemporary Models on the Target-DNA Dataset for Concentration Prediction task. Friedman ranking and Wilcoxon signed-rank test p-values (vs Our Model (Mc-GNN)) are reported for each metric. Statistically significant p-values ($p < 0.05$) are shown in **bold**.

Model	Precision			Recall			F1-Score		
	Score	**Rank**	p-val	**Score**	**Rank**	p-val	**Score**	**Rank**	p-val
Our Model	**0.9813**	**1.0**	-	**0.9818**	**1.0**	-	**0.9814**	**1.0**	-
Shiaelis *et al.* [21]	0.9484	2.0	**0.0312**	0.9494	2.0	**0.0312**	0.9475	2.0	**0.0312**
ShallowCNN-1 [11]	0.8315	3.0	**0.0312**	0.8219	3.0	**0.0312**	0.8170	3.0	**0.0312**
ShallowCNN-2 [10,22,23]	0.4104	6.8	**0.0312**	0.2857	8.0	**0.0312**	0.2104	8.0	**0.0312**
GhostNet [59]	0.5183	6.2	**0.0312**	0.4071	5.6	**0.0312**	0.3493	5.8	**0.0312**
ShuffleNet [57]	0.4768	7.2	**0.0312**	0.3705	5.8	**0.0312**	0.3356	5.0	**0.0312**
ResNet9 [61]	0.6081	5.2	**0.0312**	0.3078	7.8	**0.0312**	0.1884	8.0	**0.0312**
MobileNet [58]	0.3979	7.8	**0.0312**	0.3641	6.6	**0.0312**	0.2948	6.4	**0.0312**
EfficientNet [24,60]	0.5298	5.8	**0.0312**	0.4170	5.2	**0.0312**	0.3482	5.8	**0.0312**

The strongest baseline, Shiaelis *et al.* [21], attained the second-best rank, demonstrating the competitive power of particle-resolved fluorescence models, but with the assistance of fluorescence labels required. ShallowCNN-1 [11] followed in third place with moderate performance, highlighting the benefits and limits of lightweight 2D-convolution. In contrast, compact mobile-optimized architectures-ResNet9 [61], GhostNet [59], ShuffleNet [57], MobileNet [58], and EfficientNet [24,60], exhibited uneven precision-recall trade-offs, yielding lower F1-scores ($\leq$ 0.35) and ranks between 5.2 and 7.8. The ShallowCNN-2 [10,22,23] performed poorest, underscoring the limitations of convolutional layered image-level approaches without explicit shape or relational modelling.

As shown in Table 3, the Friedman test detected significant differences across the same evaluated baselines on the Viral-RNA concentration prediction task for all three metrics: precision ($\chi^2 = 28.53$, $p = 3.83 \times 10^{-4}$), recall ($\chi^2 = 30.88$, $p = 1.48 \times 10^{-4}$), and F1-score ($\chi^2 = 34.19$, $p = 3.80 \times 10^{-5}$). Our Mc-GNN again secured the top rank (1.0) across precision (0.9476), recall (0.9480), and F1-score (0.9470), as also presented in Table 3, with all post-hoc Wilcoxon signed-rank comparisons against baselines reaching statistical significance ($p = 0.0312$). The competitive Shiaelis *et al.* [21] again achieved the second rank (2.0), confirming its underscored value of labeled probes assisted particle-resolved fluorescence models. ShallowCNN-1 [11] ranked third, but achieved only moderate performance, confirming the limitations of simple 2D convolutional architectures in capturing complex clustering patterns. This limitation was further evidenced by reduced performance in ShallowCNN-2 [10,22,23], where increasing the number of convolutional layers led to diminishing returns. Compact architectures such as MobileNet [58]

TABLE 3 Performance Comparison of Proposed Mc-GNN with Contemporary Models on the Viral-RNA Dataset for Concentration Prediction task. Friedman ranking and Wilcoxon signed-rank test *p*-values (vs Our Model) are reported for each metric. Statistically significant *p*-values ($p < 0.05$) are shown in **bold**.

Model	Precision			Recall			F1-Score		
	Score	Rank	*p*-val	Score	Rank	*p*-val	Score	Rank	*p*-val
Our Model	**0.9476**	**1.0**	-	**0.9480**	**1.0**	-	**0.9470**	**1.0**	-
Shiaelis *et al.* [21]	0.8661	2.0	**0.0312**	0.8867	2.0	**0.0312**	0.8728	2.0	**0.0312**
ShallowCNN-1 [11]	0.5939	3.6	**0.0312**	0.5965	3.0	**0.0312**	0.5811	3.0	**0.0312**
ShallowCNN-2 [10,22,23]	0.4085	6.8	**0.0312**	0.3019	6.6	**0.0312**	0.2224	6.4	**0.0312**
GhostNet [59]	0.4302	7.2	**0.0312**	0.3319	6.2	**0.0312**	0.2073	6.8	**0.0312**
ShuffleNet [57]	0.3784	7.2	**0.0312**	0.3531	5.4	**0.0312**	0.3071	4.0	**0.0312**
ResNet9 [61]	0.4776	5.6	**0.0312**	0.3390	5.8	**0.0312**	0.1946	7.6	**0.0312**
MobileNet [58]	0.5319	5.0	**0.0312**	0.2742	7.6	**0.0312**	0.1708	7.6	**0.0312**
EfficientNet [24,60]	0.4361	6.6	**0.0312**	0.2831	7.4	**0.0312**	0.1957	6.6	**0.0312**

and EfficientNet [24,60] exhibited relatively balanced recall but suffered from low precision, resulting in F1-scores below 0.20. ShuffleNet [57] and GhostNet [59] similarly underperformed, while ResNet9 [61] achieved higher precision but failed to maintain recall and thus F1-score. These findings from both datasets underscore the effectiveness of Mc-GNN's dual-stage architecture

3 | CONCLUSIONS

The COVID-19 pandemic has underscored the global demand for rapid, affordable, and infrastructure-light POC diagnostics, a demand that conventional nucleic acid amplification and antigen tests cannot fully meet due to cost, protocol complexity, or inadequate sensitivity at low target abundance. In direct response to these limitations, we demonstrate an image-based, fluorescence-free assay that synergistically combines (i) spectrally encoded NPs that act as an optical barcode through enforced hetero-particle pairing, simultaneously bypassing the $\sim$10 nm LSPR-coupling limit and enhancing specificity by filtering out spurious aggregation, while offering a simplified, one-pot, amplification-, and enzyme-free assay design [7–9,62–64]; (ii) a simplified optical readout that enables compatibility with cost-effective, compact dark-field microscopy platforms or smartphone-based imaging [11,19], eliminating the need for bulky instrumentation and opening avenues for widespread clinical translation, and; (iii) an intuitive, morphology-guided Mc-GNN architecture for analysis at the particle level, with context-aware relational learning and feature aggregation. This integrated platform delivers laboratory-grade sensitivity (fM DNA and RNA detection, analysis across thousands of particles simultaneously) while remaining computationally efficient on consumer-grade GPUs, thereby closing the performance-accessibility gap that currently restricts widespread deployment of molecular diagnostics and eliminating the need for fluorescence or bulk colourimetric assays.

From a computational standpoint, Mc-GNN outperforms in NP image analysis. Ablation studies confirm that only the full integration of morphology-guided kernels and graph-based context fusion yields the observed performance gains. Lower performance of ShallowCNN-1 and -2 on both datasets further substantiates the need for Mc-GNN, which goes beyond conventional image-level convolution. Mc-GNN accurately classifies fM targets-98.2 % average recall for Target-DNA and 94.8 % for Viral-RNA, surpassing the particle-resolved fluorescence model of Shiaelis *et al.* and all CNN baselines (none exceeded 90% on Viral-RNA, despite the variations in NPs used and the complexity of dealing with viral materials). By uniting local morphology sensitivity with graph-level context preprocessed with particle-wise background correction and radial sorting, Mc-GNN resists imaging noise, serum interference, and vignetting, all without fluorescent labels or ROI alignment, and its design enables straightforward extension to additional analytes, NP chemistries, or multiplexed/hyperspectral modalities.

Looking ahead, the demonstrated cross-task and cross-modality robustness of Mc-GNN highlights its promise for even broader diagnostic applications. Further work can explore its extension to more diverse datasets and increased input concentration granularity that will allow Mc-GNN to be applied as a regression-based model for continuous concentration quantification. This would additionally allow a more defined limit of detection to be established, as the assay could potentially have high accuracy for concentrations below those in our current dataset. Clinical validation on real patient-derived samples, including blood or saliva, will also be critical for translation and benchmarking against standard assays, such as LFAs and RT-PCR. Notably, RT-PCR

can detect viral loads lower than 100 genome copies per mL (<1 aM)[65], which our assay currently falls seriously short of. However, our assay may achieve better results as more data points are incorporated, and it is purification-free, while RT-PCR relies on highly purified samples for best results[26,27]. Integration of additional chemical or optical modalities, such as Raman reporters or photonic barcodes, could also increase sensitivity and multiplexing capability. Finally, coupling our assay with portable microscopy and on-device AI inference could democratize molecular diagnostics, delivering high-performance testing to decentralized or resource-limited environments.

In summary, this study establishes a robust computational-biosensing framework that couples advances in NP assay design, dark-field imaging, and intuitive DL. The demonstrated performance and flexibility of Mc-GNN, particularly its ability to be applied from synthetic DNA to viral material with the same computational settings, a task of greater complexity due to the presence of serum-fortified cell growth media (which can destabilize NPs) and the requirement for viral lysis, underscore its superiority over contemporary and widely used alternatives. Notably, no other approach in our comparisons exceeded 90% accuracy on the Viral-RNA dataset, whereas Mc-GNN consistently did so, even when challenged with variations in NP selection and potential assay interferents. These results set a strong foundation for the next generation of scalable, sensitive, and accessible POC diagnostics.

METHODS

DNA Functionalization of Gold and Silver Nanoparticles

Spherical gold nanoparticles (AuNPs) (diameter, 50 nm), spherical silver nanoparticles (AgNPs) (diameter, 50 nm), and gold-shell silica-core nanoparticles (Au@SiNPs) (80 nm Si core, 20 nm Au shell) each with polyvinylpyrrolidone (PVP) (40 kDa) as capping agent were purchased from nanoComposix (USA) (Figure S2). NPs were functionalized with thiolated DNA probes (Table S3) (Integrated DNA Technologies, USA) using a modified PVP-assisted conjugation protocol[66,67]. Probe DNA sequences were chosen semi-empirically for particle stability, with initial designs created using NUPACK Web Application[68,69] and with subsequent testing at high salt after conjugation. Thiolated DNA was reduced, for use with AgNPs only, by incubation with tris(2-carboxyethyl)phosphine hydrochloride (TCEP) (Sigma Aldrich) at a 1:100 DNA:TCEP ratio for 1 hr at room temperature. DNA for use with gold was not treated, as TCEP reduction was found to be unnecessary[70]. For conjugation, PVP-capped NPs were first aliquoted from high concentration stocks (AuNP: 27.9 OD @ 528 nm, *est.* 1.2 nM; AgNP: 134.9 OD @ 419 nm, *est.* 2.7 nM; and Au@SiNP: 37.8 OD @ 666 nm, *est.* 0.15 nM) into DNA Lobind tubes (Eppendorf SE, Germany). The respective thiolated probes were added at a final concentration of 28.5 μM followed by bath sonication of the reaction (10 s)[29] and incubation (10 min) at room temperature. Additional PVP (40 kDa) (Sigma Aldrich) was added to a final concentration of 0.35% (w/v) to further stabilize the particles. A volume of salt-containing buffer, consisting of sodium phosphate buffer (PB, pH 8, 20 mM), NaBr (200 mM), and sodium dodecyl sulfate (SDS) (0. 02% (w/v), Sigma Aldrich), equivalent to the reaction volume was slowly added in a single step. Notably, NaBr was used instead of NaCl to discourage nonspecific adhesion of nucleotides to the gold/silver surface[71]. The reaction mixture was then heated at 60 °C for 2.5 hr with mixing (800 rpm). To remove residual DNA and salts, the resulting conjugated NPs were washed 5 times (or 7 times for viral detection) by centrifugation (18000*g*, 5 min) using 10 mM PB (no NaCl). Conjugated NP concentrations were then approximated using their UV-Vis Absorbance spectra (NanoDrop One Instrument, Thermofisher Scientific, USA) as compared to their reported stock concentrations.

DNA detection

A 188 bp single-stranded DNA sequence ("E-DNA") derived from the Envelope Gene and 200 bp DNA sequence ("N-DNA") derived from the Nucleocapsid Gene of SARS-CoV-2 (NCBI `NC_045512.2`) were designed for use as model DNA targets (Table S3) (Ultramer ssDNA, PAGE purified; Integrated DNA Technologies, USA). We used these two strands at equivalent concentrations to emulate a larger SARS-CoV-2 target and refer to them collectively as "Target-DNA." For Target-DNA detection, dual-probe Au@SiNPs (PrE1/PrN1-Au@SiNP) as well as two single-probe AgNP species (PrE2-AgNP, PrN2-AgNP) were used for simultaneous detection of E-DNA and N-DNA (Table S3). First, a NP mixture (in PB), a reaction buffer mixture, and Target-DNA (in TE buffer) at the relevant concentrations were prepared. The three separate solutions were mixed to start the hybridization assay. The reaction was set up such that the following final concentrations were achieved in a 40 μL volume: PB (10 mM), Au@SiNP-PrE1 (1 pM), AgNP-PrE2 (3 pM), AgNP-PrN2 (3 pM), NaCl (400 mM), Ethylenediaminetetraacetic acid

(EDTA) (2 mM, Sigma Aldrich), dextran (3% (w/v), 150 kDa, Sigma Aldrich), PVP (0.05% (w/v), 40 kDa), SDS (0.05% (w/v)), and lastly Target-DNA ([E-DNA] = [N-DNA] = reported concentration) (50 fM - 50 pM) or TE buffer as control. EDTA was included to chelate undesirable multivalent ions, while dextran was included to promote specific hybridization[72,73]. Salt and surfactant concentrations were empirically derived for optimal hybridization results (fast and specific). Immediately after mixing, the solution was vortexed and then heated to 50 °C with mixing (800 rpm) for 15 min. The reaction was allowed to cool for 10 min at room temperature before transferring 20 µL to a home-made microscopy chamber consisting of two plasma-etched coverslips sandwiching a PDMS well for optical imaging. Particles were imaged in solution using the optical setup detailed below. Notably, single Au@SiNPs settle via gravity onto the glass surface, while other particles tend to stay diffuse in solution unless they are clustered. Images were captured in an automated fashion using a piezo stage (see below) in order to generate sufficient training data for DL analysis.

Viral RNA detection

For viral RNA detection, the following reagent was deposited by the Centers for Disease Control and Prevention and obtained through BEI Resources, NIAID, NIH: SARS-Related Coronavirus 2, Isolate USA-WA1/2020, Heat Inactivated, NR-52286. Viral load was determined by BEI Resources using droplet digital PCR (stock concentration: 5.36×10^8 genome equivalents / mL). Virus was used either directly as provided or diluted in its storage medium, which consisted of Eagle's Minimal Essential Medium (EMEM) (ATCC, USA) with fetal bovine serum (FBS) (2% (v/v), Thermo Fisher Scientific, USA). Blank viral storage medium was used for virus dilution and as a control. For this assay iteration, dual-probe Au@SiNPs (PrE1/PrN1-Au@SiNP), single-probe AgNPs (PrE2-AgNP), and single-probe AuNPs (PrN2-AuNP) were used (Table S3). Therefore, the Viral-RNA could potentially cross-link all three particle species. First, a NP mixture (in PB), a reaction (lysis) buffer mixture, and the target/control (in viral storage medium) at the relevant concentrations were prepared. The three separate solutions were mixed to start the viral detection assay. The reaction was set up such that the following final concentrations were achieved in a 20 µL volume: PB (10 mM), Au@SiNP-PrE1/PrN1 (1 pM, AgNP-PrE2 (3 pM), AuNP-PrN2 (3 pM), NaCl (200 mM), EDTA (50 mM), dextran sulfate (3% (w/v), 500 kDa, Sigma Aldrich), SDS (0.5% (w/v)), Triton-X100 (0.5%, Sigma Aldrich), Proteinase K (10 mg/mL, Thermo Fisher Scientific, USA) and lastly inactivated SARS-CoV-2 (10 µL) at the relevant dilution (or viral storage medium as control). SDS, Triton-X100, and Proteinase K were included at relevant concentrations for viral lysis and nuclease inhibition[74]. Upon mixing, the sample was treated as described above for heating and imaging.

Optical imaging

Images, including spectral information, of different NP sub-populations was collected using a colour (RGB) camera (Teledyne FLIR Grasshopper 3, model GS3-U3-89S6C-C, Sony IMX255 sensor). The camera was integrated into a custom-built inverted dark-field microscope that also housed broad-spectrum light-emitting-diodes (LEDs) (Cree C503D-WAN, 40k mcd intensity) as an affordable light source and a microscope objective (Olympus LUCPlanFL N 20x 0.45) to image the samples. Subtle spectral variations are classified using a supervised DL approach, which automates the counting and differentiation of NP sub-populations. To generate labeled training data for the DL model, we implemented an imaging protocol to systematically capture images of particles in an automated fashion using piezo stages (Thorlabs PD2/M 5 mm Linear Stage with Piezoelectric Inertia Drive).

Raw Image Preprocessing and Dataset Preparation

Each raw microscopy image is subjected to a systematic preprocessing steps aimed at isolating individual NP clusters as well as unbound particles, which are often visually indistinguishable. This process is illustrated in Figure 2.

In preparation for particle-level feature extraction, raw images were subjected to a preprocessing pipeline aimed at accurate and uniform particle segmentation. First, each RGB image was converted to grayscale to simplify subsequent operations. Contrast enhancement was applied using CLAHE[53], which improves local contrast and emphasizes fine-grained structures by limiting noise amplification in homogeneous regions. To reduce high-frequency noise and improve contour clarity, a Gaussian blur was then applied. Using this preprocessed image, a preliminary mask of particles was created. This mask was then subtracted from the original image to estimate the background region. The average background intensity was computed and subsequently subtracted from the raw image to correct for uneven illumination and staining artifacts. Following background correction, manual

thresholding was applied to the grayscale image to facilitate precise contour detection. Contours corresponding to individual particles were then sorted based on the Euclidean distance of their centroids from the center of the image. This prioritization is biologically and optically motivated: particles closer to the optical axis tend to be better focused, more centrally illuminated, and less affected by vignetting or edge distortion. This spatial sorting ensures that particle quality remains consistent across the dataset and minimizes inter-sample variability due to imaging conditions. For each detected particle, a slightly enlarged region surrounding its contour was defined to estimate the local background specific to that particle. A binary mask of the particle was created within this region, and an exclusion mask was generated to avoid interference from nearby particles. The clean background region-free from both the particle and neighbors-was used to compute an accurate local background intensity, which was subtracted from the corresponding particle region to enhance signal fidelity. The resulting particle region was then cropped to isolate the individual particle.

Each extracted particle was resized to a fixed dimension of 15×15×3 (height × width × RGB channels) using inter-area interpolation, which is well-suited for downsampling as it preserves edge fidelity and reduces aliasing. Particles smaller than the target size were padded with zeros to maintain a uniform shape. All particles from a given image were then converted into tensors and stacked along the batch dimension, forming a data tensor of shape [N, 3, 15, 15] for downstream parallel processing. The stacked particle representation serves as the input to the proposed Mc-GNN and baseline models for particle-wise feature extraction. Although features are extracted for all particles in each raw image, a mean aggregation across the particle dimension N yields a final image-level feature embedding of dimension [1280]. This aggregated representation is then passed through a classifier comprising two fully connected layers. For concentration prediction task, the output layer uses a number of neurons equal to the number of classes.

Due to the limited number of high-quality images remaining after filtering out unfocused and blurry samples, particularly since the test set typically comprised only ∼20% of the data, performing 5-fold cross-validation was essential to rigorously assess model generalizability across the entire dataset. Stratified cross-validation was employed to ensure that the distribution of concentration classes remained consistent across all folds, thereby minimizing sampling bias and allowing fair comparison of model performance under uniform class proportions. All models were thus trained and evaluated using 5-fold stratified cross-validation. In each fold, 80% of the raw images were allocated for training and validation (subdivided into 75% training and 25% validation), while the remaining 20% were reserved for testing. The experiments were implemented in PyTorch 2.4.1 with CUDA 12.1 and executed on an NVIDIA RTX A5000 GPU. Each training was performed for 300 epochs using the Adam optimizer with a learning rate of 1×10^{-5}, with early stopping based on validation accuracy using a patience of 50 epochs.

Proposed architecture and Implementation

To accurately capture particle-level spectral, spatial and morphological information in NP-based assays for concentration prediction, we propose Mc-GNN, a hybrid convolutional-graph architecture (Figure 3).

Input and Morphology-guided Kernel Design. As illustrated in Figure 2, following the input raw image preprocessing and the dataset preparation steps, the input image is processed as batched tensor $\mathbf{X} \in \mathbb{R}^{N\times 3\times 15\times 15}$, where N denotes the number of particle images in a batch. To embed domain-specific inductive biases into the learning process, we introduce a set of shape-aware masked convolutional kernels. Specifically, five structural masks-*circular, edge, Gaussian, donut, and square*-are utilized to modulate the learnable kernels. Each of these masks emphasizes distinct structural properties relevant to NP morphology: **Circular (C)**: emphasizes isotropic particle intensity, **Gaussian (G)**: captures center-focused (diffuse or gradient-like structures), smooth intensity distributions, **Donut (D)**: highlights ring-like particle structures, **Square (S)**: captures rectilinear or grid-like features, **Edge (E)**: enhances particle boundaries. Randomly initialized 252 kernels $\mathbf{K} \in \mathbb{R}^{252\times 3\times 7\times 7}$ are element-wise modulated by 252 of each of these 5 masks $\mathbf{M}^{(s)}$, producing morphology-guided kernels. The selection of 252 kernels per shape was empirically determined via non-extensive hyperparameter tuning to optimize model performance. After normalization, convolution of these morphology-guided kernels with input particle tensors directly encodes morphological priors into the resulting feature maps, making them explicitly shape-aware at the earliest feature extraction stage:

$$\mathbf{F}^{(s)} = \mathrm{Conv2D}(\mathbf{X}, \mathbf{K}^{(s)}, \mathrm{stride} = 3, \mathrm{padding} = 3).$$

Spatial global average pooling compresses feature maps into compact vectors. The pooled outputs from all five shape branches to form are stacked to form a unified tensor of shape $[N, 5, 252]$, which is then permuted into the shape $[N * 5, 252]$ to facilitate downstream graph-based relational modelling.

Graph-based relational modelling and GCN based relational learning. The spatially compressed and permuted tensor $\widetilde{\mathbf{F}} \in \mathbb{R}^{[N\times 5,252]}$ encodes, for each particle, five parallel morphology-guided features, each as a 252-dimensional embedding derived from a distinct morphology-guided kernel branch. This multi-view representation serves as a foundation for capturing context-aware relational learning across morphology-guided kernel outputs. Thus, corresponding to each input particle, we represent the above tensor as a graph with 5 nodes, where each is a 252-dimensional morphology-guided embedding vector. The inter-relations between these morphology-guided kernel outputs for each particle are captured via a learnable adjacency matrix $\mathbf{A} \in \mathbb{R}^{5\times 5}$, where each entry A_{ij} denotes the importance of information flow from node j to node i.

$$\mathbf{H}^{(\ell+1)} = \sigma\left(\hat{\mathbf{A}}\mathbf{H}^{(\ell)}\mathbf{W}^{(\ell)}\right),$$

where $\hat{\mathbf{A}}$ is a normalized adjacency matrix learned to represent kernel interactions, $\mathbf{W}^{(\ell)}$ are learnable weights, and $\sigma(\cdot)$ denotes a ReLU activation.

Two GCN layers act on the graph representation of morphology-guided particle-feature embeddings to produce final particle embeddings $\mathbf{z} \in \mathbb{R}^{[252*5=1280]}$ upon flattening node representations. A GCN layer aggregates and transforms features from a node's neighbors in a graph, enabling context-aware relational reasoning by capturing structural dependencies and shared context across connected nodes. Thus, this graph-based relational modeling passed through GCN layers enables Mc-GNN to dynamically capture complementary interactions among morphology-guided kernel outputs. By leveraging the unique strengths of each kernel type through GCN-based relational learning, the architecture adapts to diverse binding patterns and concentration conditions-enhancing interpretability and discriminative power in concentration prediction for the NP-based assays.

Aggregation and Classification. As shown in Figure 2, each raw image yields thousands of cropped particles (typically N = 1,000-5,000), for which morphology-guided and graph-enhanced features are extracted in parallel by Mc-GNN. These particle-wise embeddings are then mean-aggregated to produce a single [1280]-dimensional raw input image-level representation, operated further by fully connected layers for the final quantification. The output layer adapts to the task: multiple neurons equal to the number of classes for concentration prediction.

Deep Learning Evaluation Protocol

Model performance was quantitatively assessed using standard classification metrics, including Precision, macro Recall (Recall), and F1-score. Macro Recall is computed by first calculating the recall for each class independently, and then averaging these values across all classes. This treats each class equally, regardless of its frequency in the dataset, and is therefore particularly appropriate for imbalanced classification problems. In the context of our concentration prediction task, macro Recall is equivalent to average class-wise accuracy, as it reflects the model's ability to correctly identify each concentration class on a per-class basis before averaging. F1-score, the harmonic mean of Precision and Recall, was also computed for each class and then averaged (macro F1-score) to provide a balanced measure of performance that accounts for both false positives and false negatives. This is particularly useful when evaluating multiclass classification under class imbalance, where precision-recall tradeoffs may vary across classes. For all metrics, scores were computed on the held-out test set for each fold in a 5-fold stratified cross-validation scheme, and the final reported values represent the mean performance aggregated across all folds. Due to the multiclass and imbalanced nature of the concentration prediction task, AUC was not reported for this setting.

To assess the statistical significance of observed performance differences among models, we employed a rigorous non-parametric evaluation protocol. The Friedman test was first applied to the performance scores of all models across the cross-validation folds, with ranks computed to summarize each model's relative performance. The Friedman test evaluates whether the observed ranking differences across models are likely to have occurred by chance; statistical significance is established if the test yields a low p-value (e.g., $p < 0.05$) and a large χ^2 statistic, indicating that at least some models perform differently from others. Upon detecting significant overall differences, we conducted Wilcoxon signed-rank post-hoc tests[56] for robust pairwise comparison between the proposed Mc-GNN and each baseline. These tests were performed on the fold-wise metric distributions obtained from identical data splits for all models. This non-parametric framework not only mitigates the impact of outliers and non-Gaussian metric distributions, but also minimizes bias arising from sampling variation, thereby providing a statistically robust foundation for comparative claims.

DECLARATIONS

SUPPORTING INFORMATION

Additional results and analysis for the analyte detection task, including confusion matrices, detailed performance comparison tables (Target-DNA and Viral-RNA datasets). Includes UV-Vis absorbance spectra and agarose gel electrophoresis characterization of nanoparticles with distinct scattering properties, and complete DNA and probe sequences used in the assays.

ACKNOWLEDGEMENTS

We thank Professor Atul N. Parikh (NTU - Singapore and UC Davis) for valuable discussions and for reviewing the final manuscript. Figure 1a contains art adapted from Servier Medical Art (https://smart.servier.com/), licensed under CC BY 4.0 (https://creativecommons.org/licenses/by/4.0/).

FUNDING

This research was funded by the Ministry of Education (MOE), Singapore, through a Center of Excellence Grant to the Institute for Digital Molecular Analytics and Science (IDMxS, grant: EDUN C-33-18-279-V12) at Nanyang Technological University.

CONFLICT OF INTEREST

The authors declare no potential conflict of interests.

DATA AND CODE AVAILABILITY

The data that support the findings of this study are available from the corresponding author(s) upon reasonable request.

AUTHOR CONTRIBUTION

K.J., X-L.C., and J.B.D. contributed equally to this work. K.J., X-L.C., J.B.D., P.T. and L.W. performed conceptualization. X-L.C., J.B.D., L.A.P., S.B., and F.N.N.M.H. performed experimental investigation. K.J. developed the data processing and deep learning method, carried out model training, and analyzed the results. Y.Y. and Y.W. contributed to model-related discussions. P.T. and L.W. performed supervision. K.J., X-L.C., and J.B.D. wrote the original manuscript draft. All authors contributed to review and editing of the final manuscript.

References

1. Rezaei M, Bazaz SR, Zhand S, et al. Point of Care Diagnostics in the Age of COVID-19. *Diagnostics.* 2021;11(1):9. doi: 10.3390/diagnostics11010009

2. Valera E, Jankelow A, Lim J, et al. COVID-19 Point-of-Care Diagnostics: Present and Future. *ACS Nano.* 2021;15(5):7899–7906. doi: 10.1021/acsnano.1c02981

3. Posthuma-Trumpie GA, Korf J, Amerongen vA. Lateral flow (immuno)assay: its strengths, weaknesses, opportunities and threats. A literature survey. *Analytical and Bioanalytical Chemistry.* 2009;393(2):569–582. doi: 10.1007/s00216-008-2287-2

4. Mistry DA, Wang JY, Moeser ME, Starkey T, Lee LYW. A systematic review of the sensitivity and specificity of lateral flow devices in the detection of SARS-CoV-2. *BMC Infectious Diseases.* 2021;21(1):828.

5. Mirkin CA, Letsinger RL, Mucic RC, Storhoff JJ. A DNA-based method for rationally assembling nanoparticles into macroscopic materials. *Nature.* 1996;382(6592):607–609. doi: 10.1038/382607a0

6. Alivisatos AP, Johnsson KP, Peng X, et al. Organization of 'nanocrystal molecules' using DNA. *Nature.* 1996;382:609–611.

7. Kim S, Park JE, Hwang W, et al. Optokinetically encoded nanoprobe-based multiplexing strategy for microRNA profiling. *Journal of the American Chemical Society.* 2017;139(9):3558–3566.

8. Seo J, Kim S, Park HH, Choi DY, Nam JM. Nano-bio-computing lipid nanotablet. *Science advances.* 2019;5(2):eaau2124.

9. Kim S, Kim N, Seo J, et al. Nanoparticle-based computing architecture for nanoparticle neural networks. *Science Advances.* 2020;6(35):eabb3348.

10. Tun WST, Talodthaisong C, Daduang S, et al. A machine learning colorimetric biosensor based on acetylcholinesterase and silver nanoparticles for the detection of dichlorvos pesticides. *Materials Chemistry Frontiers.* 2022;6(11):1487–1498.

11. Sajed S, Kolahdouz M, Sadeghi MA, Razavi SF. High-Performance Estimation of Lead Ion Concentration Using Smartphone-Based Colorimetric Analysis and a Machine Learning Approach. *ACS Omega.* 2020;5:27675–27684.

12. Li X, Yang X, Pan Z, Zhuo S, Lin Z, Chen J. Nucleic acid amplification-free biosensor for sensitive and specific cfDNA detection based on CRISPR-Cas12a and single nanoparticle dark-field microscopy (DFM) imaging. *Sensors and Actuators: B. Chemical.* 2024;406:135363.

13. Zhang D, Wang K, Wei W, Liu Y, Liu S. Multifunctional Plasmonic Core-Satellites Nanoprobe for Cancer Diagnosis and Therapy Based on a Cascade Reaction Induced by MicroRNA. *Anal. Chem..* 2021;93:9521–9530.

14. Bennett D, Chen X, Walker GJ, et al. Machine Learning Color Feature Analysis of a High Throughput Nanoparticle Conjugate Sensing Assay. *Anal. Chem..* 2023;95:6550–6558.

15. Orouji A, Ghasemi F, Hormozi-Nezhad MR. Machine Learning-Assisted Colorimetric Assay Based on Au@Ag Nanorods for Chromium Speciation. *Anal. Chem..* 2023;95:10110–10118.

16. Wu N, Zhang XY, Xia J, Li X, Yang T, Wang JH. Ratiometric 3D DNA Machine Combined with Machine Learning Algorithm for Ultrasensitive and High-Precision Screening of Early Urinary Diseases. *ACS Nano.* 2021;15:19522–19534.

17. Yang J, Li G, Chen S, et al. Machine Learning-Assistant Colorimetric Sensor Arrays for Intelligent and Rapid Diagnosis of Urinary Tract Infection. *ACS Sens..* 2024;9:1945–1956.

18. Li B, Li X, Dong Y, et al. Colorimetric Sensor Array Based on Gold Nanoparticles with Diverse Surface Charges for Microorganisms Identification. *Anal. Chem..* 2017;89:10639–10643.

19. Song C, Zhou L, Wang Y, et al. Deep Learning-Assisted Ultrasensitive Detection of Gold Nanoparticles Using Light Microscopy Images Captured by a Cellphone Camera. *Anal. Chem..* 2025;97:5164–5170.

20. Wang XY, Hong Q, Zhou ZR, Jin ZY, Li DW, Qian RC. Holistic Prediction of AuNP Aggregation in Diverse Aqueous Suspensions Based on Machine Vision and Dark-Field Scattering Imaging. *Anal. Chem..* 2024;96:1506–1514.

21. Shiaelis N, Tometzki A, Peto L, et al. Virus Detection and Identification in Minutes Using Single-Particle Imaging and Deep Learning. *ACS Nano.* 2023;17:697–710.

22. Liu S, Zhao J, Wu J, et al. A deep learning-enabled smartphone platform for rapid and sensitive colorimetric detection of dimethoate pesticide. *Analytical and Bioanalytical Chemistry.* 2023;415(29):7127–7138.

23. Kang J, Yoo YJ, Park JH, et al. DeepGT: Deep learning-based quantification of nanosized bioparticles in bright-field micrographs of Gires-Tournois biosensor. *Nano Today.* 2023;52:101968.

24. Jin KX, Shen J, Wang YJ, Yang Y, Cao SH. Enhanced Nanoparticle Recognition via Deep Learning-Accelerated Plasmonic Sensing. *Biosensors.* 2024;14(8):363.

25. McAffee DB, Hu Q, Arnob A, Wu HJ, Groves JT. Machine-learning strategy for high error tolerance in image-based digital molecular assays. *bioRxiv.* 2025:2025–08.

26. Watzinger F, Ebner K, Lion T. Detection and monitoring of virus infections by real-time PCR. *Molecular Aspects of Medicine.* 2006;27(2):254-298.

27. Chu DKW, Pan Y, Cheng SMS, et al. Molecular Diagnosis of a Novel Coronavirus (2019-nCoV) Causing an Outbreak of Pneumonia. *Clinical Chemistry.* 2020;66(4):549–555.

28. Bevilacqua P, Nuzzo S, Torino E, Condorelli G, Salvatore M, Grimaldi AM. Antifouling Strategies of Nanoparticles for Diagnostic and Therapeutic Application: A Systematic Review of the Literature. *Nanomaterials.* 2021;11(3).

29. Hurst SJ, Lytton-Jean AKR, Mirkin CA. Maximizing DNA Loading on a Range of Gold Nanoparticle Sizes. *Anal. Chem..* 2006;78(24):8313–8318.

30. Patterson EI, Prince T, Anderson ER, et al. Methods of Inactivation of SARS-CoV-2 for Downstream Biological Assays. *The Journal of Infectious Diseases.* 2020;222(9):1462–1467.

31. Cao J, Leng H, Lischinski D, Cohen-Or D, Tu C, Li Y. Shapeconv: Shape-aware convolutional layer for indoor rgb-d semantic segmentation. In: Proceedings of the IEEE/CVF international conference on computer vision. 2021:7088–7097.

32. Tek FB, Cam I, Karlı D. Adaptive convolution kernel for artificial neural networks. *Journal of Visual Communication and Image Representation.* 2021;75:103015.

33. Li X, Wang W, Hu X, Yang J. Selective kernel networks. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2019:510–519.

34. Chen LC, Papandreou G, Kokkinos I, Murphy K, Yuille AL. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. *IEEE transactions on pattern analysis and machine intelligence.* 2017;40(4):834–848.

35. Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2015:3431–3440.

36. Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. In: Springer. 2015:234–241.

37. Wang J, Sun K, Cheng T, et al. Deep high-resolution representation learning for visual recognition. *IEEE transactions on pattern analysis and machine intelligence.* 2020;43(10):3349–3364.

38. Zhang P, Zhu J, Zhang C, Rong Z, Wu Y. Semi-global shape-aware attention network for image segmentation and retrieval. *Neurocomputing.* 2022;506:369–379.

39. Kuo W, Angelova A, Malik J, Lin TY. Shapemask: Learning to segment novel objects by refining shape priors. In: Proceedings of the IEEE/CVF international conference on computer vision. 2019:9207–9216.

40. Liu S, De Mello S, Gu J, Zhong G, Yang MH, Kautz J. Learning affinity via spatial propagation networks. *Advances in Neural Information Processing Systems.* 2017;30.

41. Takikawa T, Acuna D, Jampani V, Fidler S. Gated-scnn: Gated shape cnns for semantic segmentation. In: Proceedings of the IEEE/CVF international conference on computer vision. 2019:5229–5238.

42. Liu X, Tian B, Wang Z, et al. Delving into shape-aware zero-shot semantic segmentation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2023:2999–3009.

43. Feng A, You C, Wang S, Tassiulas L. Kergnns: Interpretable graph neural networks with graph kernels. In: . 36. Proceedings of the AAAI conference on artificial intelligence. 2022:6614–6622.

44. Zhang L, Li X, Arnab A, Yang K, Tong Y, Torr PH. Dual graph convolutional network for semantic segmentation. *arXiv preprint arXiv:1909.06121.* 2019.

45. Xu J, Dai E, Luo D, Zhang X, Wang S. Shape-aware graph spectral learning. In: Proceedings of the 33rd ACM International Conference on Information and Knowledge Management. 2024:2692–2701.

46. Bera A, Wharton Z, Liu Y, Bessis N, Behera A. SR-GNN: Spatial relation-aware graph neural network for fine-grained image categorization. *IEEE Transactions on Image Processing.* 2022;31:6017–6031.

47. Zhou Y, Graham S, Alemi Koohbanani N, Shaban M, Heng PA, Rajpoot N. Cgc-net: Cell graph convolutional network for grading of colorectal cancer histology images. In: Proceedings of the IEEE/CVF international conference on computer vision workshops. 2019:0–0.

48. Zhu T, Yao G, Hu D, Xie C, Gong H, Li A. Morphognn: morphological embedding for single neuron with graph neural networks. *bioRxiv.* 2022:2022–05.

49. Shen Y, Zhou B, Xiong X, Gao R, Wang YG. How graph neural networks enhance convolutional neural networks towards mining the topological structures from histology. In: . 8. ICML Workshop on Computational Biology. 2022.

50. Scarselli F, Gori M, Tsoi AC, Hagenbuchner M, Monfardini G. The Graph Neural Network Model. *IEEE Transactions on Neural Networks.* 2009;20(1):61-80. doi: 10.1109/TNN.2008.2005605

51. Jeon Y, Kim J. Active convolution: Learning the shape of convolution for image classification. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2017:4201–4209.

52. Dai J, Qi H, Xiong Y, et al. Deformable convolutional networks. In: Proceedings of the IEEE international conference on computer vision. 2017:764–773.

53. Zuiderveld K. *Contrast limited adaptive histogram equalization*:474–485; USA: Academic Press Professional, Inc. . 1994.

54. Pawley J. *Handbook of biological confocal microscopy*. 236. Springer Science & Business Media, 2006.

55. Goldman DB. Vignette and Exposure Calibration and Compensation. *IEEE Transactions on Pattern Analysis and Machine Intelligence.* 2010;32(12):2276-2288. doi: 10.1109/TPAMI.2010.55

56. Demšar J. Statistical comparisons of classifiers over multiple data sets. *Journal of Machine learning research.* 2006;7(Jan):1–30.

57. Zhang X, Zhou X, Lin M, Sun J. Shufflenet: An extremely efficient convolutional neural network for mobile devices. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2018:6848–6856.

58. Howard AG, Zhu M, Chen B, et al. Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv preprint arXiv:1704.04861.* 2017.

59. Han K, Wang Y, Tian Q, Guo J, Xu C, Xu C. Ghostnet: More features from cheap operations. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2020:1580–1589.

60. Tan M, Le Q. Efficientnet: Rethinking model scaling for convolutional neural networks. In: Proceedings of Machine Learning Research (PMLR). 2019:6105–6114.

61. Moddy2024 . ResNet-9. https://github.com/Moddy2024/ResNet-9; 2024. Accessed: 2025-06-19.

62. Sönnichsen C, Reinhard BM, Liphardt J, Alivisatos AP. A molecular ruler based on plasmon coupling of single gold and silver nanoparticles. *Nature Biotechnology.* 2005;23:741–745.

63. Chen JIL, Chen Y, Ginger DS. Plasmonic Nanoparticle Dimers for Optical Sensing of DNA in Complex Media. *J. Am. Chem. Soc.*. 2010;132:9600–9601.

64. Pennacchio A, Giampaolo F, Cafaro V, et al. A bacterial biosensor based on gold nanoparticles functionalized by a hydrophobin-chimera and combined with machine learning for user-friendly detection. *Sensors and Actuators: B. Chemical.* 2024;410:135645.

65. Pickering S, Batra R, Merrick B, et al. Comparative performance of SARS-CoV-2 lateral flow antigen tests and association with detection of infectious virus in clinical specimens: a single-centre laboratory evaluation study. *The Lancet Microbe.* 2021;2(9):e461–e471.

66. Heo JH, Kim KI, Cho HH, et al. Ultrastable-Stealth Large Gold Nanoparticles with DNA Directed Biological Functionality. *Langmuir.* 2015;31(51):13773–13782.

67. Xu S, Yuan H, Xu A, Wang J, Wu L. Rapid Synthesis of Stable and Functional Conjugates of DNA/Gold Nanoparticles Mediated by Tween 80. *Langmuir.* 2011;27(22):13629–13634.

68. Zadeh JN, Steenberg CD, Bois JS, et al. NUPACK: Analysis and design of nucleic acid systems. *Journal of Computational Chemistry.* 2011;32:170–173.

69. Fornace ME, Huang J, Newman CT, Porubsky NJ, Pierce MB, Pierce NA. NUPACK: Analysis and Design of Nucleic Acid Structures, Devices, and Systems. *ChemRxiv.* 2022;Version 1.

70. Wu R, Jiang LP, Zhu JJ, Liu J. Effects of Small Molecules on DNA Adsorption by Gold Nanoparticles and a Case Study of Tris(2-carboxyethyl)phosphine (TCEP). *Langmuir.* 2019;35(41):13461–13468.

71. Liu B, Wu P, Huang Z, Ma L, Liu J. Bromide as a Robust Backfiller on Gold for Precise Control of DNA Conformation and High Stability of Spherical Nucleic Acids. *J. Am. Chem. Soc.*. 2018;140(13):4499–4502.

72. Amasino RM. Acceleration of nucleic acid hybridization rate by polyethylene glycol. *Analytical Biochemistry.* 1986;152(2):304-307.

73. Donnelly T, Faulds K, Graham D. Investigation of Silver Nanoparticle Assembly Following Hybridization with Different Lengths of DNA. *Particle & Particle Systems Characterization.* 2016;33:404–411.

74. Bender AT, Sullivan BP, Lillis L, Posner JD. Enzymatic and Chemical-Based Methods to Inactivate Endogenous Blood Ribonucleases for Nucleic Acid Diagnostics. *The Journal of Molecular Diagnostics.* 2020;22(8):1030–1040.

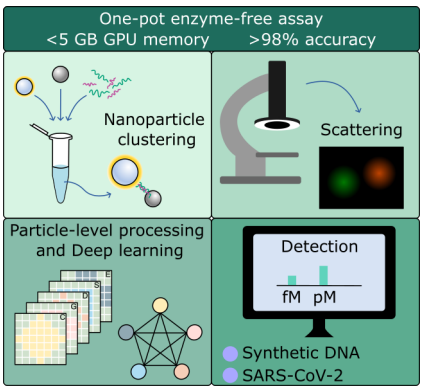

FIGURE 4 For Table of Contents (ToC) Only: An affordable, one-pot diagnostic test that achieves femtomolar sensitivity for nucleic acids without enzymatic amplification is presented. The assay induces target-induced clustering of nanoparticles, imaged via a low-cost dark-field imaging system. A novel morphology-guided convolutional graph neural network analyzes thousands of particles per image achieving 98.2% accuracy with synthetic DNA and 94.8% with SARS-CoV-2 RNA, outperforming contemporary methods on low-memory GPUs.