Nanoveu Limited
ACN 624 421 085
Level 45, 108 St Georges Terrace
Perth WA, 6000 Australia
+61 8 6244 9095
www.nanoveu.com

ASX RELEASE **12 November 2026**

ASX: NVU

Nanoveu's EMASS Delivers Up to 56× Lower Energy Consumption than Leading Published Competitors on Purpose-Built AI Models

EMASS unveils new ECS-DoT-22 benchmark results showing the lowest published energy per inference on every MLPerf® Tiny workload, with sub-microjoule inference, higher accuracy than the reference models and up to 9× faster processing on the same silicon

For personal use only

Highlights

- **Up to 56× lower energy consumption:** with models built for its accelerator for the first time, ECS-DoT-22 now sits below every published MLPerf Tiny result on energy per inference for keyword spotting, visual wake words, image classification and anomaly detection. The new models use 17× to 56× less energy per inference than the leading commercial platforms with published results on each workload (the MLCommons-verified Syntiant Core 2 on three workloads, the BrainChip AKD1000 on anomaly detection).
- **Higher accuracy and up to 9× speedup on the same silicon:** the new models, built for the ECS-DoT deep-learning accelerator, exceed the MLPerf Tiny reference models on every official test set, clear the quality thresholds by up to 6.3 percentage points, and at 50 MHz run 9× faster on keyword spotting, 4.5× on image classification and 3.0× on anomaly detection.
- **Sub-microjoule inference:** 0.56 µJ per inference for keyword spotting and 0.39 µJ for anomaly detection. At ten keyword-spotting inferences a second the AI compute cost is approximately 6 µW, less than a thousandth of what a Bluetooth or Wi-Fi radio uses to send the same audio to the cloud, which is why the AI belongs on the sensor.
- **Major addressable markets:** most sensors, wearables, hearables and industrial monitors run on a battery or harvested energy and today either stream raw data to the cloud or run no AI at all. At microjoules per decision, ECS-DoT-22 lets them run always-on AI on the sensor itself and transmit data only when there is something worth sending.
- **Replicable results:** following the DigiKey distribution agreement announced on 10 November 2026, ECS-DoT-22 evaluation kits can be ordered off the shelf. With the EMASS IDE and AI compiler due for release this month, customers and independent third parties will be able to run the same benchmark models and scripts and reproduce the results.
- **Positioned for customer Design-In opportunities:** the results provide technical evidence to support engagement with semiconductor customers and original equipment manufacturers evaluating ultra-low-power edge AI solutions. Commercial adoption remains subject to customer qualification, integration and procurement decisions.

Cautionary statement: *ECS-DoT-22 benchmark results presented are Company-measured and have not been independently verified or certified by MLCommons. Comparisons use identified published results from different benchmark rounds and include both verified and unverified peer results. They do not establish superiority over every semiconductor currently available.*

Nanoveu Limited (ASX: NVU, OTCQB: NNVUF) ("Nanoveu" or the "Company"), a technology company specialising in advanced semiconductor, drone and materials sciences, is pleased to announce new benchmark results for ECS-DoT-22, the 22nm ultra-low-power edge-AI system-on-chip developed by its wholly owned subsidiary Embedded A.I. Systems Pte Ltd ("EMASS"), on all four MLPerf Tiny workloads: keyword spotting, visual wake words, image classification and anomaly detection.

ECS-DoT is built for always-on, battery-powered devices, where the energy to make a decision needs to be lower than the energy to transmit the data wirelessly (e.g., using Bluetooth). MLPerf Tiny, maintained by MLCommons, is the industry benchmark for AI inference on microcontroller-class devices. For each workload it specifies a reference neural network, a dataset and a minimum accuracy, and it measures latency and energy per inference. Nanoveu first reported ECS-DoT results on the reference networks in October and November 2024. Those results placed ECS-DoT-22 up to 20× below the Syntiant Core 2 on energy per inference, and the Company has used that comparison in its investor material since.

The results announced today are the first on models optimised for the ECS-DoT deep-learning accelerator ("DLA") rather than for general-purpose microcontrollers. Each new model had to equal or exceed the reference model's accuracy on the same official test set before its latency and energy were measured. All four new models did. On the same chip at 50 MHz they achieve up to 9× speedup than the reference models and, on three of the four workloads, use less energy per inference. EMASS reports both sets of results, so that ECS-DoT-22 can be compared with any peer on the same basis that peers reported.

ECS-DoT-22 Energy per Inference

Workload	New model	Reference model (2024)	Change
Keyword spotting	0.564 µJ	3.07 µJ	5.4× lower
Image classification	2.67 µJ	5.5 µJ	2.1× lower
Anomaly detection	0.385 µJ	0.8 µJ	2.1× lower
Visual wake words	3.94 µJ	3.7 µJ	6% higher

Table 1. New-model energy per inference measured on the evaluation board at 50 MHz with a 0.66 V core supply. Reference-model figures are those published in October and November 2024.

Keyword spotting and anomaly detection now run below one microjoule per inference. At ten keyword-spotting inferences a second, the energy for AI compute is approximately 6 µW; the same duty at the Syntiant Core 2's lowest-energy operating point would require approximately 315 µW. Visual wake words is the exception: the new model is more accurate and marginally faster than the reference, at 6% more energy.

For personal use only

For personal use only

ECS-DoT-22 Accuracy and Latency

Workload (dataset)	Accuracy, new model	Accuracy, reference model	MLPerf Tiny threshold	Latency, new model	Latency, reference model	Speed-up
Keyword spotting (Speech Commands v0.02)	92.13%	91.66%	90%	434 µs	3.9 ms	9.0×
Image classification (CIFAR-10)	88.25%	86.97%	85%	1.392 ms	6.3 ms	4.5×
Anomaly detection (ToyADMOS ToyCar)	AUC 0.861	AUC 0.8595	AUC 0.85	397 µs	1.2 ms	3.0×
Visual wake words (COCO 2014, 96×96)	86.30%	84.80%	80%	3.474 ms	5.2 ms	1.5×

Table 2. Measured on ECS-DoT-22 at 50 MHz. Accuracy is over the full official test set, run on the evaluation board. Latency is one inference on the chip, from input in memory to result. Reference-model latencies are those published on 22 October and 12 November 2024.

Against Published Peer Results

Figure 1 shows energy per inference for ECS-DoT-22 against the published MLPerf Tiny results from active vendors that report both latency and energy for the same workload: the MLCommons-verified submissions of Asygn, STMicroelectronics, Syntiant, Silicon Labs and Renesas, and the unverified results published by Ambiq and BrainChip. Where a vendor submitted more than one system or operating point, its lowest-energy result is shown. Table 3 sets each ECS-DoT-22 result against the lowest published result on that workload, the Asygn ColibryNPU, and against the Syntiant Core 2 and BrainChip AKD1000 results the Company has reported since 2024, and Figure 2 plots latency against energy for the same set.

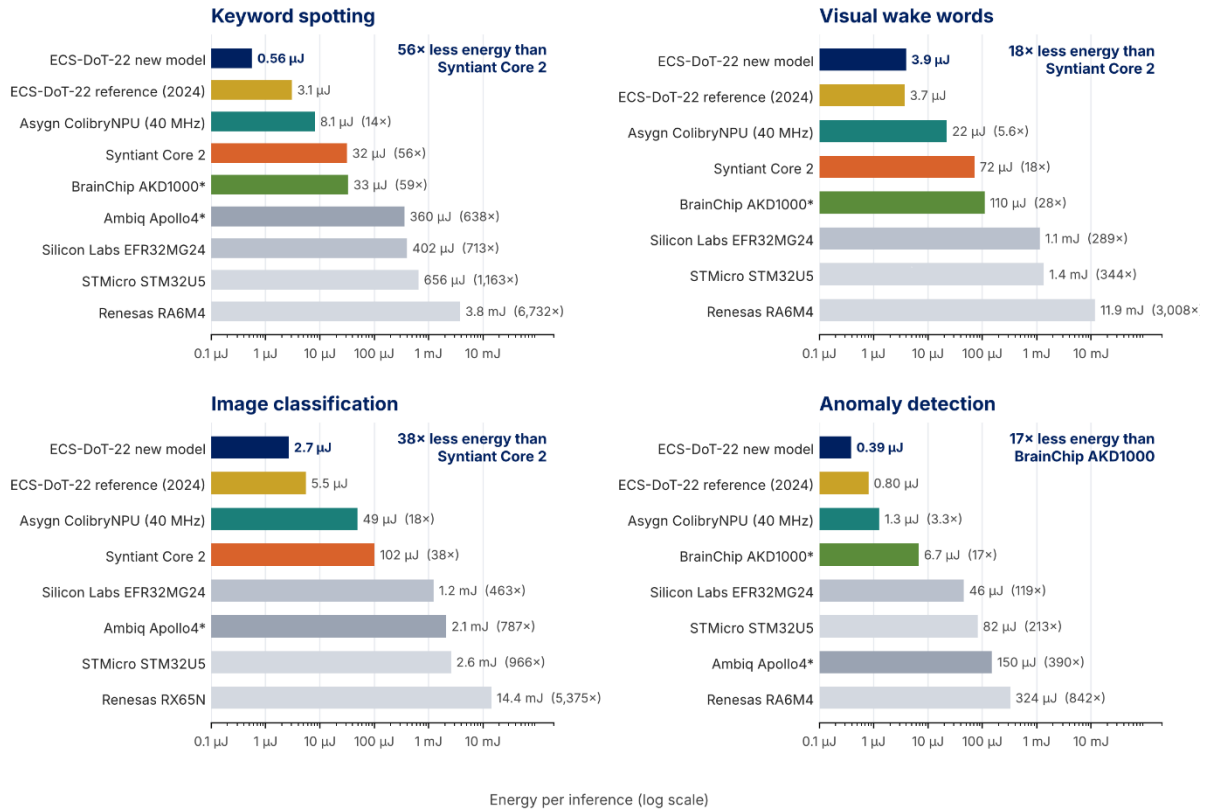


Figure 1: Energy per inference by workload, ECS-DoT-22 against published MLPerf Tiny results

Workload	ECS-DoT-22, new model	Lowest published result: Asygn ColibryNPU, 40 MHz (verified)	Energy/ EDP benefits	Syntiant Core 2 (verified) / BrainChip AKD1000 (unverified)	Energy/ EDP advantage	Latency advantage
Keyword spotting (KWS)	0.434 ms · 0.564 µJ	4.26 ms · 8.07 µJ	**14×/140x**	Syntiant Core 2: 4.4 ms · 31.5 µJ	56×/560x	10×
Visual wake words (VWW)	3.47 ms · 3.94 µJ	12.3 ms · 22.2 µJ	**5.6×/20x**	Syntiant Core 2: 12.7 ms · 71.7 µJ	18×/66.6x	3.7×
Image classification (IC)	1.39 ms · 2.67 µJ	20.2 ms · 49.2 µJ	**18×/261x**	Syntiant Core 2: 16.0 ms · 101.8 µJ	38×/418x	11×
Anomaly detection (AD)	0.397 ms · 0.385 µJ	0.433 ms · 1.27 µJ	**3.3×/3.6x**	BrainChip AKD1000: 0.46 ms · 6.73 µJ	17×/20.4x	1.2×

Table 3. Lowest published result is the lowest energy per inference filed with MLCommons by any active vendor on each workload: Asygn ColibryNPU P1, MLPerf Tiny v1.4 Closed, 40 MHz operating point, verified by MLCommons Association, retrieved from mlcommons.org on 12 November 2026; at its 300 MHz point the ColibryNPU records 0.567 ms · 17.8 µJ, 1.63 ms · 49.6 µJ, 2.69 ms · 108 µJ and 0.058 ms · 2.74 µJ, and is faster than ECS-DoT-22 on visual wake words and anomaly detection. Syntiant Core 2 (NDP120): MLPerf Tiny v1.0 to v1.2 Closed, 0.9 V / 30.7 MHz operating point, verified by MLCommons Association; BrainChip AKD1000: unverified result published by BrainChip, January 2023; these are the comparisons the Company has reported since 2024. Energy and latency advantage are the peer's figure divided by ECS-DoT-22's; ECS-DoT-22 is also faster than the ColibryNPU at 40 MHz on all four workloads (9.8×, 3.5×, 14.5× and 1.1×, following the same order in the table). Energy-delay product (energy per inference multiplied by inference time) benefits: against the ColibryNPU 140×, 20×, 261× and 3.6× (highlighted in Table 3) at 40 MHz and 41×, 5.9×, 78× and 1.0× at 300 MHz; against Syntiant and BrainChip at their lowest-energy operating points 560×, 67×, 440× and 20×, and against Syntiant's 1.1 V / 98.7 MHz point 270×, 29× and 190× (KWS, VWW, IC). The lowest-energy MLCommons-verified anomaly-detection result from an active vendor other than Asygn is Silicon Labs' EFR32MG24 at 45.7 µJ, 119× ECS-DoT-22's. ECS-DoT-22: Company-measured.

On all four workloads the new models are faster than the Syntiant Core 2 and the BrainChip AKD1000 and use 17× to 56× less energy per inference. Taking the two together, the energy-delay product, energy per inference multiplied by latency, is 20× to 560× lower than the Syntiant Core 2's and BrainChip AKD1000's at their lowest-energy operating points, and 20× to 270× lower at Syntiant's fastest. The Syntiant energy comparison is the most conservative available: it uses the lowest-energy of Syntiant's two submitted operating points, 0.9 V and 30.7 MHz, rather than the 1.1 V and 98.7 MHz point at which the Core 2 runs faster but uses more energy; both points are shown in Figure 2.

For personal use only

For personal use only

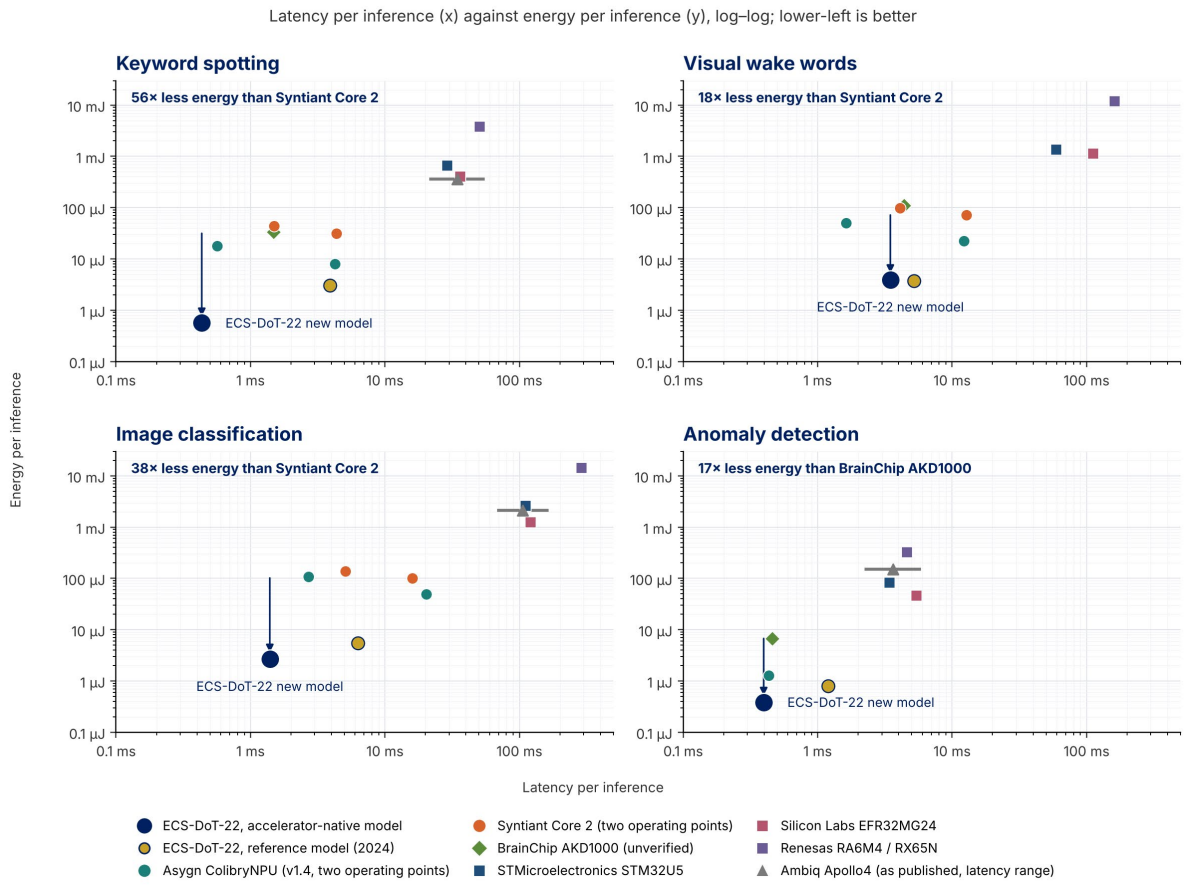


Figure 2: Latency against energy per inference, log–log; both Syntiant and ColibryNPU operating points shown

On the unmodified reference models, the results the Company reported in 2024, ECS-DoT-22 already used 10× to 19× less energy than the Syntiant Core 2. The accelerator-native models widen a lead the reference models already held. Against the Apollo4 Plus figures Ambiq published in October 2022, the new models use approximately 400× to 800× less energy per inference on keyword spotting, image classification and anomaly detection.

Why EMASS Reports Two Sets of Results

MLPerf Tiny's closed division fixes the network, so that silicon can be compared on identical models. Its open division allows the model to change, provided accuracy is measured on the same official test set and every departure is documented. Vendors publishing outside MLCommons, including BrainChip and Analog Devices, report results on networks shaped for their own accelerators, and no product team ships the reference networks, which were written for general-purpose microcontrollers. The accelerator-native model is what a customer will receive from the EMASS IDE. EMASS therefore reports both: the unmodified reference models, for a like-for-like comparison, and its own models held to at least the reference accuracy. The reference-model results in this release follow the closed-division model. The accelerator-native results are open-division-style results.

What the Results Mean in a Product

- **The AI drops out of the power budget.** At ten keyword-spotting inferences a second the compute cost is approximately 6 µW. Always-on listening or watching no longer sets the size of the battery.
- **The accelerator finishes early and sleeps.** Sub-millisecond keyword spotting and anomaly detection leave the accelerator idle for most of each second, with headroom to run several models on one sensor frame.
- **Fewer false triggers.** Higher accuracy than the reference on every workload means fewer missed wake words and fewer false person detections, which a user notices before battery life.

Who Publishes Both Latency and Energy

Among active vendors, five have filed both latency and energy results with MLCommons: Asygn, STMicroelectronics, Syntiant, Silicon Labs and Renesas. Ambiq, BrainChip and EMASS publish both outside MLCommons. Syntiant has no anomaly-detection result and BrainChip none for image classification.

Company	Product	Keyword spotting	Visual wake words	Image classification	Anomaly detection	Where published
EMASS	ECS-DoT-22	Y	Y	Y	Y	ASX releases 2024; this release
STMicroelectronics	STM32 (L4, U5, H7, U3)	Y	Y	Y	Y	MLCommons v0.7–v1.4
Syntiant	NDP120 / Core 2	Y	Y	Y	N	MLCommons v0.7–v1.4
Silicon Labs	EFR32MG24	Y	Y	Y	Y	MLCommons v0.7, v1.0
Renesas	RA6M4, RX65N	Y	Y	Y	Y	MLCommons v0.7
Ambiq	Apollo4 Plus	Y	Y	Y	Y	Ambiq, October 2022
BrainChip	Akida AKD1000	Y	Y	N	Y	BrainChip white paper, January 2023
Asygn	ColibryNPU	Y	Y	Y	Y	MLCommons v1.4

Table 4. Y = a published result with both latency and energy for that workload in any round. MLCommons rows are taken from the per-submitter results folders in mlcommons/tiny_results v0.5 to v1.4; a Y requires both a performance and an energy result. Software-only submitters (Krai, OctoML, Skymizer) and companies that filed latency only (Qualcomm, Nuvoton, Andes, Bosch) are not shown.

For personal use only

How the Results Were Measured

Iso-accuracy. Each new model must reach at least the reference model's accuracy on the same official test set. Its latency and energy are then compared with the reference model's on the same chip.

Test sets. Speech Commands v0.02 test set (4,890 clips, 12 classes); the 1,000 EEMBC visual-wake-words images; CIFAR-10 test set (10,000 images); 248 ToyCar evaluation files. Feature extraction follows MLPerf Tiny's own code. No test data is used in training.

Energy. Energy per inference is derived from the current drawn by the whole chip from its core supply rail during one inference, measured on the ECS-DoT-22 evaluation board at 50 MHz and a 0.66 V core supply. Other board-level components are not included.

Availability. ECS-DoT-22 evaluation kits are in customers' hands and can be ordered through DigiKey. The EMASS integrated development environment with its AI compiler is being developed for release in November 2026, and the benchmark models, projects and host scripts behind every table in this release will be made available with it, so that customers and independent third parties can repeat the tests.

Dr Mohamed Sabry, Director of Nanoveu and Founder of EMASS, said: *"Until now every ECS-DoT benchmark used networks written for microcontrollers. These models were built for our accelerator, and the difference shows: more accurate than the reference on every workload, up to nine times faster, and keyword spotting and anomaly detection below a microjoule. This is the class-leading performance the next generation of always-on devices needs, across hearables, wearables, smart devices, industrial sensing, robotics and drones, and with the evaluation kits now on DigiKey, it is performance customers can put on their own bench."*

Mark Goranson, Nanoveu's CEO of Semiconductor Technologies, said: *"Every company at Design-In stage runs ECS-DoT-22 against its own workload before it commits, and the first question is always energy per inference. These results give our sales team and our customers' engineers one set of numbers, measured on the chip they are evaluating, against every published competitor. The margin is wide enough to change what a product team thinks is possible on a battery."*

Notes

The ECS-DoT-22 results presented in this announcement are company-measured results and have not been independently verified by MLCommons Association or subjected to its formal MLPerf® review process. The reference-model results follow a methodology based on the MLPerf Tiny Closed Division, while the accelerator-native results use modified models consistent with an Open Division-style approach. Neither set constitutes an official MLCommons-verified submission.

Unverified benchmark results may use measurement methodologies, model architectures or workload implementations that differ from those prescribed for verified MLPerf results. Accordingly, comparisons between verified and unverified results may not be directly equivalent.

ECS-DoT-22 energy per inference is derived from the current drawn by the whole chip from its core supply rail during inference, at 50 MHz and a 0.66 V core supply for the new models. Other supply rails and board-level components are not included, and the figures do not represent total device or system-level energy consumption.

Comparative benchmark results are based on identified publicly available results from the relevant manufacturers and MLCommons benchmark submissions, as described in this announcement. They do not necessarily represent the latest or highest-performing semiconductor products available from each manufacturer.

Peer results verified by MLCommons Association: MLPerf® Tiny v1.0 Closed (Silicon Labs), v1.0 to v1.2 Closed (Syntiant), v1.2 Closed (STMicroelectronics), v1.4 Closed (Asygn), v0.7 Closed (Renesas and Silicon Labs anomaly detection), retrieved from mlcommons.org on 5 and 12 November 2026. Ambiq (October 2022) and BrainChip (January 2023) figures are unverified results published by those companies.

The MLPerf name and logo are trademarks of MLCommons Association in the United States and other countries. All rights reserved. Unauthorized use strictly prohibited.

For personal use only

Sources: Banbury et al., MLPerf Tiny Benchmark (arXiv 2106.07597); MLCommons tiny_results repositories v0.5 to v1.4 (github.com/mlcommons); Syntiant, Core 2 MLPerf Tiny v1.1 results (July 2023); Ambiq, Apollo4 Plus benchmark results (October 2022); BrainChip, Benchmarking AI Inference at the Edge (January 2023); Asygn, ColibryNPU MLPerf Tiny v1.4 results (July 2026); Nanoveu ASX announcements of 22 October 2024, 12 November 2024, 12 October 2026 and 10 November 2026.

This announcement has been authorised for release by the Board of Directors.

Further information:

Nanoveu Media

E: info@nanoveu.com

About Nanoveu Limited

Further details on the Company can be found at https://nanoveu.com/.

EMASS is a fabless semiconductor company specialising in ultra-low-power, AI-enabled system-on-chip (SoC) solutions for always-on sensing and intelligent processing at the edge. Its ECS-DoT architecture enables real-time AI inference at milliwatt-level power, supporting next-generation applications across drones, wearables, hearables, smart devices, industrial IoT, robotics and autonomous systems.

Spinoff Robotics is a Singapore-based aerial robotics company that designs and builds proprietary drone platforms. Its two current platforms are ALICE, a tethered aerial system engineered for high-payload, extended-duration operations with resistance to GPS and radio-frequency jamming and spoofing, and METRON, a camera-based 3D measurement system for confined-space inspection, delivering millimetre-level precision with AI-driven deviation detection. Both platforms have been validated in commercial deployments with Tier-1 Singaporean customers. Spinoff Robotics provides drone and sensing capabilities complementing the Company's ECS-DoT edge-AI silicon across defence, critical-infrastructure security and industrial inspection markets.

EyeFly3D™ is a comprehensive platform solution for delivering glasses-free 3D experiences across a range of devices and industries. At its core, EyeFly3D™ combines advanced screen technology, sophisticated software for content processing, and the potential for integration with ultra-low-power SoC hardware.

Nanoshield™ includes protective and functional coating technologies designed for use across a range of applications, including solar and other industrial surfaces.

Forward-Looking Statements

This announcement contains 'forward-looking information' that is based on the Company's expectations, estimates and projections as of the date on which the statements were made. This forward-looking information includes, among other things, statements with respect to the Company's business strategy, plans, development, objectives, performance, outlook, growth, cash flow, projections, targets and expectations and related expenses. Generally, this forward-looking information can be identified by the use of forward-looking terminology such as 'outlook', 'ambition', 'anticipate', 'project', 'target', 'potential', 'likely', 'believe', 'estimate', 'expect', 'intend', 'may', 'mission', 'would', 'could', 'should', 'scheduled', 'will', 'plan', 'forecast', 'evolve' and similar expressions. Persons reading this announcement are cautioned that such statements are only predictions, and that the Company's actual future results or performance may be materially different. Forward-looking information is subject to known and unknown risks, uncertainties and other factors that may cause the Company's actual results, level of activity, performance, or achievements to be materially different from those expressed or implied by such forward-looking information. Readers should not place undue reliance on forward-looking statements. The Company does not undertake to update or revise any forward-looking statement.

For personal use only