Healthcare Artificial Intelligence Value Proposition: A White Paper

Abstract

According to the Wall Street Journal, deep learning systems have a fatal flaw: there is no way to tell how a decision was made. CytoSavvy is a pathology informatics software company using artificial intelligence to attack the big data problem for drug discovery and personalized medicine with a novel shape detection methodology packaged as a browser-based solution. CytoSavvy's patented Shape Based Modeling Segmentation (SBMS) software collects critical data not available from other sources and stores this data in a relational database. We are targeting high volume drug discovery and cancer grading applications currently performed 100% manually. Our competitive differentiators are transparency and reliability. This white paper describes three critical criteria necessary to successfully employ artificial intelligence and deep learning systems.

Contact:

John T. Freyhof
Chief Executive Officer
CytoSavvy
611 Vassar Road
Wayne, PA 19087
610-225-0498 X251 (Ph)
610-225-0499 (Fax)
john@cytosavvy.com
www.cytosavvy.com

Healthcare Artificial Intelligence Value Proposition

The Opportunity

According to a recent Frost & Sullivan market analysis on the Healthcare Artificial Intelligence Solutions Market, the addressable healthcare analytics market size for Cytosavvy products has the potential to reach $6.7 billion by 2021 (growing at a staggering 40% CAGR) *source: AlleyWatch*. Recent CytoSavvy business development activities and funding for several high profile deep learning companies (PathAI $11 million, Prognos $30 million, and Zebra $12 million) confirms this. CytoSavvy's proprietary technologies provide a powerful platform to pursue a new line of high margin artificial intelligence data mining opportunities in multiple medical verticals.

The Problem

Academic medical centers, hospitals, insurance companies, big pharma, and commercial diagnostic Labs are under pressure to increase efficiency, improve quality, improve patient safety, and outcomes. Personalized medicine, drug discovery, and information management are critical to attacking these problems along with a strong need to manage huge datasets and convert that data into meaningful and actionable insights. Moore's law is helpful, but not enough. A new paradigm, particularly in digital imagery, is needed to provide highly reliable answers as good as or better than a human.

The CytoSavvy Solution

CytoSavvy has developed an algorithm that uses shaped-based modeling segmentation (SBMS) to reduce the variability identifying cancer versus non-cancer in digital pathology images (See Attachments 1, 2, and 3). Our software is designed to analyze large images and deliver a clinically relevant decision in a SQL database. This is accomplished utilizing scale and angle invariant Bezier curves. Effectively, we are taking image-analysis out of the realm of pixels and into more reliable mathematical models. The SBMS platform has game-changing potential for collecting data from digital imagery and can significantly improve the accuracy of deep learning systems by providing a powerful set of tools to address overfitting and the bias-variance tradeoff (which frequently result in accuracy rates of less than 60% vs. over 95% for CytoSavvy). SBMS adds a powerful platform to extract hard data from digital imagery and significantly enhances solutions for predictive modeling and natural language processing (see **Fractals and Disease: A New Paradigm** below)

The Unmet Need

In September 2017, the FDA decided to allow digital imagery from whole slide scanners to become a primary diagnostic tool in addition to glass slides and frozen tissue specimens. This decision has created a large-scale data mining problem for insurance companies, hospitals, and big pharma. Right now, millions of digital images are not analyzed because of a lack of

appropriate tools. SBMS provides a critical bridge to the production of hard data associated with these images that can be linked with electronic medical records and test results.

Artificial intelligence and deep learning systems used for image analysis have a fatal flaw: there is no way to explain how their decisions are made. Cancer diagnostics and drug discovery involve life and death decisions. We, as drug makers, physicians, and service providers are accountable for how those decisions are made: whether to a patient, an insurance company, or a jury

CytoSavvy's patented image analysis methodology produces a pictorial record of how its diagnostic or research decisions are made in addition to producing hard data that can be displayed in a spreadsheet. More importantly, it provides a powerful methodology to refine those decisions by through a set fully auditable rules. We call this methodology Shape Based Modeling Segmentation

CytoSavvy's patented SBMS technology makes it possible to capture robust fully automated whole side analysis data in an industry standard SQL database to attack the following currently unsolved problems:

- Understanding the data: 2D capture of 3D information
- Ingesting images from different scanners with widely varying quality and brightness levels without retraining
- Integration of image information
    - Cellular
    - Molecular
    - Phenotypic
    - Genomic
- Measuring tumor heterogeneity in an objective way

According to the Digital Pathology Association what is needed now is:

- An image analysis algorithm intended for primary diagnosis (e.g., for H&E stains)
- Whole tissue-based image analysis algorithms, as opposed to the Region of Interest/Field-of-View based algorithms that currently have clearance
- Functionality that significantly reduces the role of the pathologist in interpreting images or rendering a diagnosis while still allowing the pathologist to make the final decision.

Fractals and Disease: A New Paradigm

Despite many years of *research*, a method to precisely and quantitatively determine cancer disease state remains elusive, until now. The National Institutes of Health through the *National*

Center for Biological Information and others have suggested using *fractal dimensionality* as a metric for disease state. This is because as cells become more diseased their fractal dimension decreases because they are recapitulating earlier embryological conditions. SBMS is the key to unlocking the door to this linkage.

Current practice for characterizing solid tumors involves the use of varying systems of tumor grading and staging and thus leaves diagnosis and clinical staging dependent on the experience and skill of the practitioner. Although numerous disease markers have been identified, no combination of them has been found that produces a quantifiable and reliable measure of disease state. Newly developed genomic markers (and other measures based on the developing sciences of complexity) offer promise that this situation may soon be changed for the better by combining SBMS data with other available information: cellular, molecular, phenotypic, and genomic.

Background

Vy DBA CytoSavvy was recently awarded an Air Force Research Lab Contract to package our Shape-Based Modeling Segmentation Platform (SBMS) on a chip to decrease processing time and reduce weight requirements for unmanned aerial systems (UAS) countermeasures (swarm attacks and a moving camera motion detector for people, vehicles, and weapons). Screen captures of our first generation moving camera motion detector are outlined in Figure 1. Our program manager is encouraging us to partner with a larger organization to move things along more quickly and we are dialoguing with a senior product manager at Intel.

Figure1: Moving Camera Motion Detector

In 2014 we pivoted and re-purposed our military technology towards personalized medicine and drug discovery under the brand name CytoSavvy. CytoSavvy has generated a lot of traction in the oncology space to use (SBMS) to reduce the variability identifying cancer versus non-cancer in digital pathology images.

CytoSavvy's value proposition is time savings (typically eighty percent), consistency, risk mitigation, quality management, and improved outcomes for manual processes involving large digital pathology images. Our key competitive differentiators are reliability (typically over ninety percent) and the ability to collect hard image analysis data in a SQL database or spreadsheet (transparency and auditability).

Our Revenue Model encompasses per image processing fees (typically in the $5 – 50 per slide range depending upon the application), custom development at $125/Hour, licensing fees (typically in excess of $150,000), and annual maintenance fees

Our software is designed to analyze very large images and deliver a clinically relevant decision in a SQL database. This is accomplished utilizing scale- and angle-invariant Bezier curves. Effectively, we are taking image analysis out of the realm of pixels and into more reliable mathematical models.

Two separate modes of reliability enhancement are afforded by this approach:

- First, By fitting mathematical models of smooth curves to large numbers of pixels we diminish the effect of single pixel errors
- Second, the geometrical smoothing imposed by the mathematical models greatly decreases the number of false positives, thereby eliminating large amounts of spurious data that would have to be processed elsewhere.

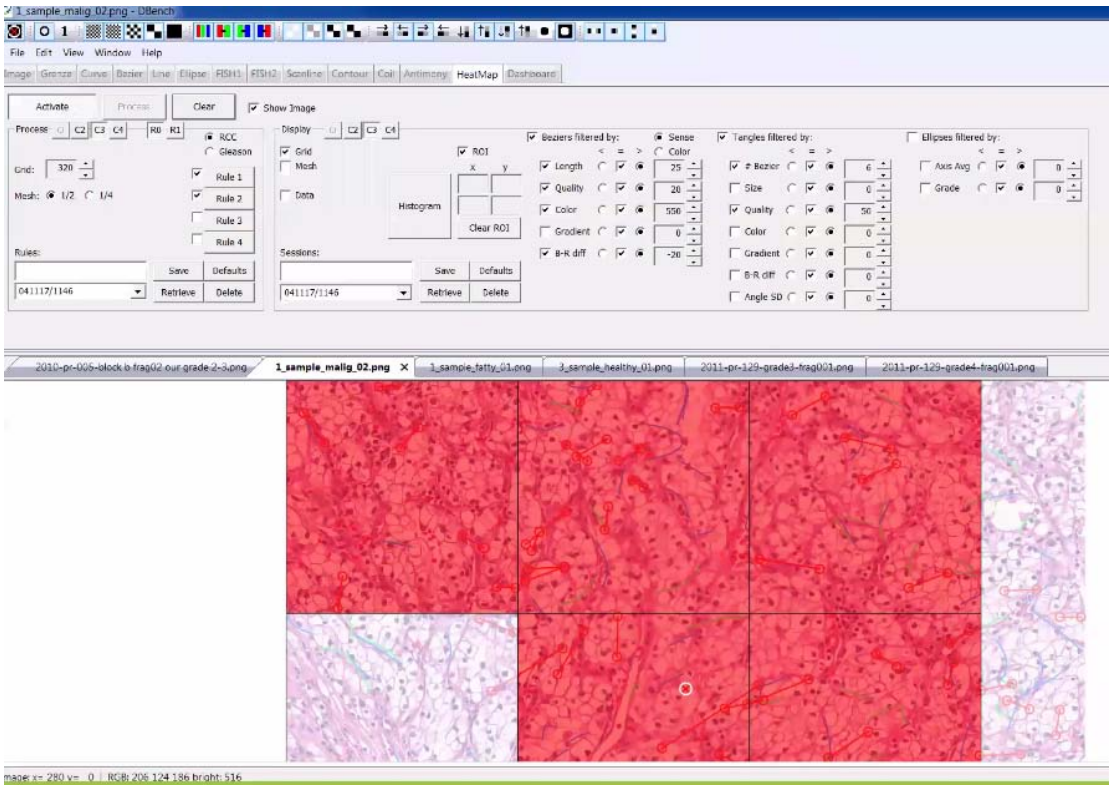


Figure 2: dBench Shape-Based Modeling Segmentation Platform

SBMS can significantly improve deep learning algorithm results by adding a shape-based filter for qualifying texture and photometric results. See Figure 2: dBench SBMS Development Platform. For UAS applications, Bezier curves offer massive compression (particularly for on-board computing). Our platform has lots of potential uses, but for now, we have tried to stay focused packaging tumor grading as a stand-alone browser based application for research projects or integrated as a module in a larger system.

The dBench SBMS development platform provides a powerful set of pre-packaged shape primitives to extract data from digital pathology images. It encompasses rule sets for tumor and tissue heterogeneity (areas of interest), as well as classifying and characterizing cells. Platform results can be packaged as a stand-alone browser based application for research projects or integrated as a module in a larger system. A sample of the data elements is included as Figure 3:

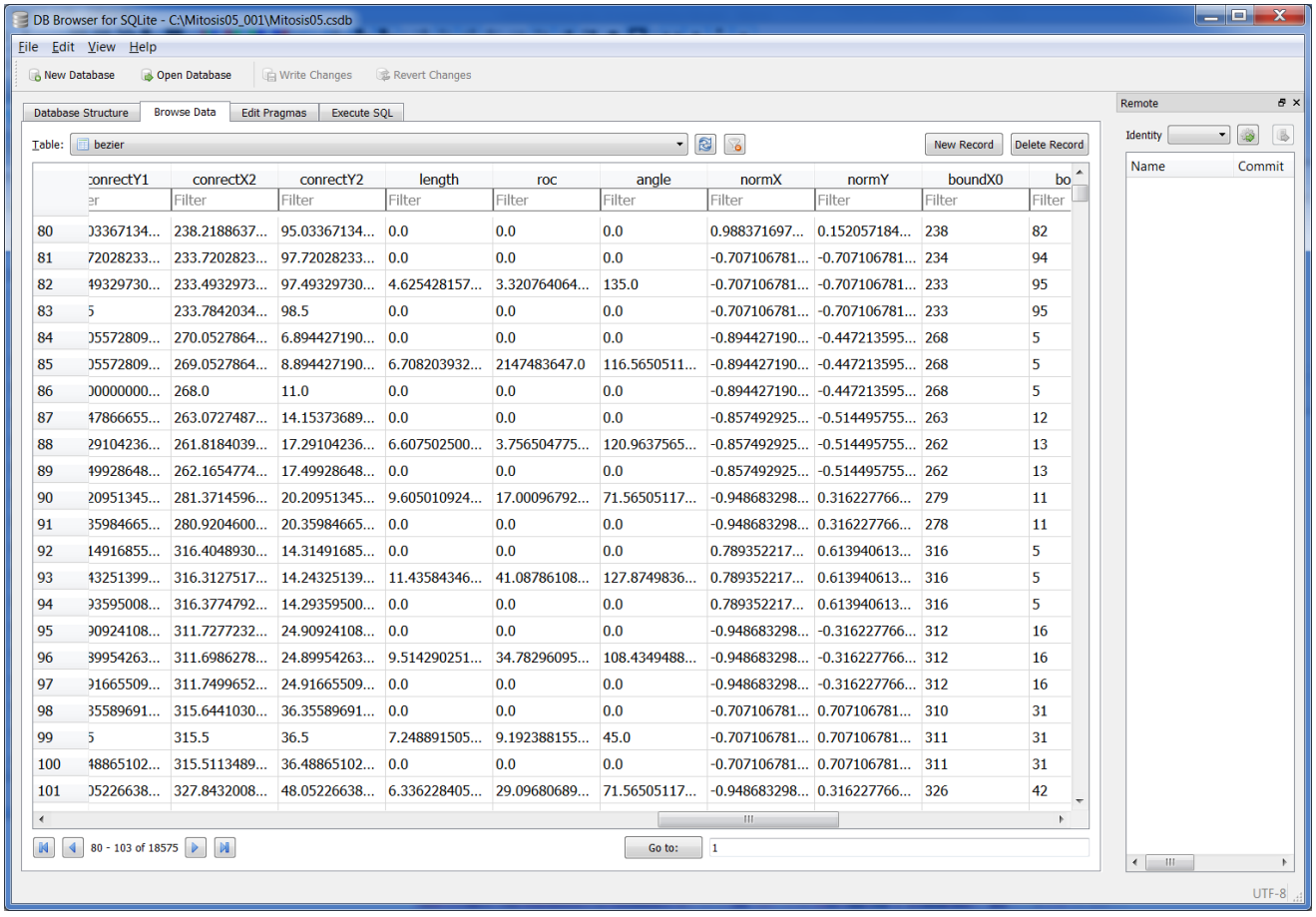


Figure 3: SQL Image Data

Recommendations and Next Steps

Artificial intelligence and deep learning systems used for image analysis have a fatal flaw: there is no way to explain how their decisions are made. If you, or any of the teams you are responsible for, are using or contemplating using a deep learning system, you need to make sure you can answer "yes" to the following three questions:

1. Can you visualize the result in an image?
2. Can the results of your decision be displayed in a spreadsheet?
3. Is there a feedback mechanism to understand how the decision was made?

If the answer to any of these questions is no, *you need to look for a better alternative.* This problem has been identified as a serious systemic flaw by industry experts and regulators; (Wall Street Journal, Wired, TechRepublic)

Attachment 1: Shape Based Modeling Segmentation Differentiators

SBMS begins with preliminary edge detection, but uses a far more robust approach to it. Edges are assessed with multiple threshold calculations (MT), then edges are further assessed at different binned pixel resolutions (multiple resolutions), resulting in vastly improved accuracy beyond typical simplistic edge approaches. While many approaches stop here, this is only the first step for SBMS.

Feature/Description	SBMS	Traditional Segmentation	Exhibit Reference
Original Image	Yes	Yes	One
Edge Detection	Yes	Yes	Two
Bezier curves derived from MTMR edges (first shape detection step)	Yes	No	Three
Bezier Edge Elements (selects one or more point on each Bezier that corresponds to a structural element of a nucleus, uses search trees to find edges)	Yes	No	Four
Associate each Edge Element with Exactly One Nucleus (uses search trees to associate groups of edges within one nucleus)	Yes	No	Five
Discover perimeter edges for each nucleus	Yes	Sometimes	Six
Fit ellipse to perimeter elements (reliably identify location, size, establish eliptical coordinate system to search for detail within the nucleus)	Yes	No	Six

Figure 1: SBMS Segmentation Versus Traditional Methods

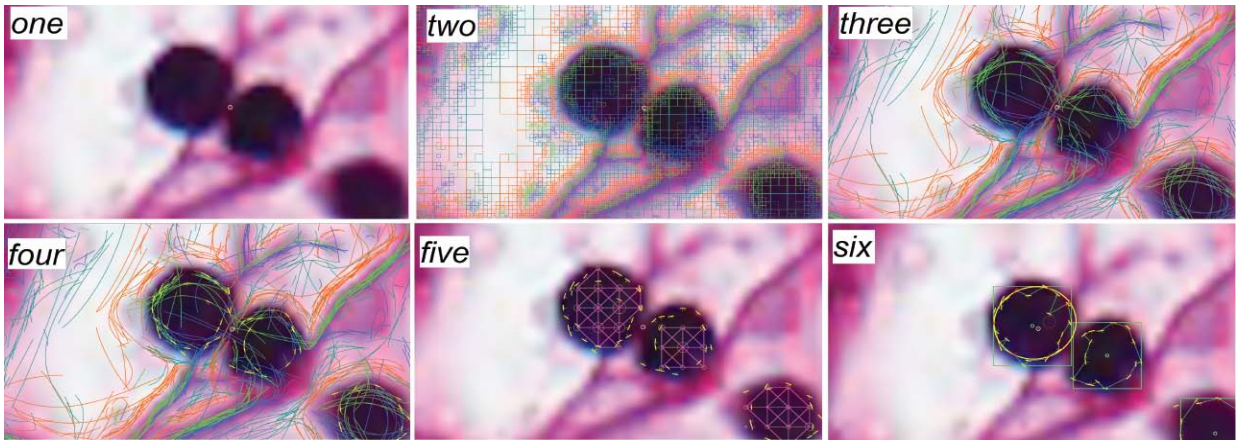
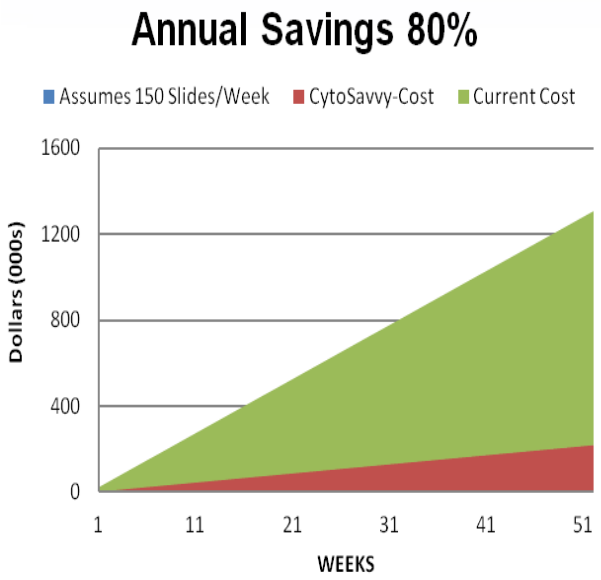


Figure 2: SBMS Methodology

Figure 2 describes the SBMS methodology applied to nucleus segmentation, as viewed in the CytoSavvy development Workbench. **One:** Portion of a raw image displaying two closely aligned nuclei. **Two:** Edge detection after identifying color and intensity gradients in image. **Three:** Bezier curves derived from edge data. **Four:** Decision trees are applied, identifying one or more points from many Bezier curves that best fit nuclei. **Five:** Decision trees associate each edge element with exactly one nucleus, establishing a grid pattern on each object to allow further intra-object examination. **Six:** reliably establish perimeter of nuclei and fit ellipse to determine, size, location, and interior areas of interest. Significance: the system finds 95% of targeted nuclei with less than .1 % false positives.

Attachment 2: Cost and Time Savings Per Slide

Outlined below are the expected savings using CytoSavvy in a typical pathology lab utilizing digital imagery for renal cell carcinoma tumor grading. Payback ranges from three to twelve months depending upon which edition or subscription plan most closely matches expected usage volume.



Benefits:

- No Workflow Change: works with all major microscope vendors
- Browser-Based Interactive Graphics are Easy to Use with minimal training
- Eighty Percent Faster than Manual Analysis
- Improves Outcomes by Providing Pro Forma Decision for Approval

Savings per slide at a typical institution average eighty percent for our targeted applications. This means that a busy pathologist can typically accomplish in about twelve minutes what used to take sixty minutes with better data and complete documentation for our targeted applications. Current automation techniques can only achieve accuracy rates in the zero to forty percent range. CytoSavvy's cell counting and characterization technology is typically finding over 95% of targeted nuclei with less than .1 percent false positives. A table outlining a typical savings scenario follows.

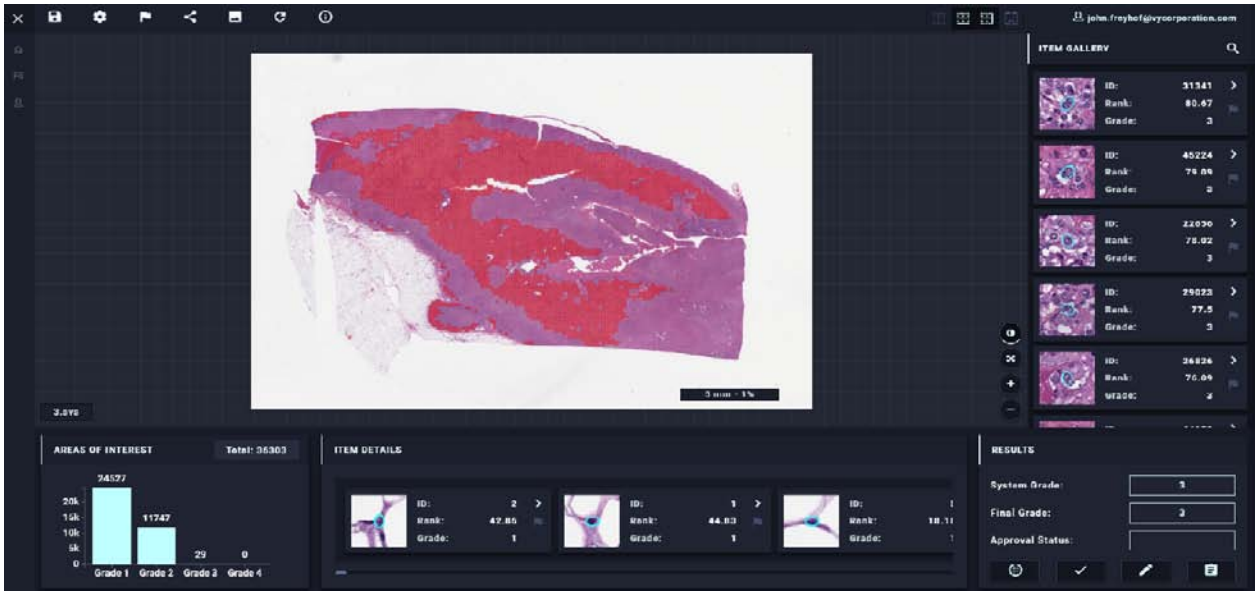
	Current			Cytosavvy		
	Time(2)	Cost(3)	Total	Time	Cost	Total
Sample Preparation	0.2	$50	$10		$50	$0
Slide Scan	0.1	$10	$1		$10	$0
Analytics						
* Automated	0.2	$5	$1	0.1	$5	$1
* Manual	0.4	$75	$30	0.2	$75	$15
Diagnostics	0.5	$75	$38	0.1	$75	$8
Data Collection	0.8	$75	$60	0.07	$75	$5
Total Cost			$140			$28
Data Value Index (1)			5			8

Based upon an informal survey of lab technicians at three institutions.
On a scale of 1 - 10, how does researcher value the data collected by the system

Assumes 60 minutes per hour. .1 = 6 minutes, .2 = 12 minutes

Cost assumes that Manual slide analysis, diagnostics, and data collection are conducted by an MD Pathologist or Ph.D.

Attachment 3: Browser Based Dashboard With Native Image, Analytics Row, and Gallery Column



Kidney (Renal Cell Carcinoma) Dashboard

The Kidney dashboard features the native image in the upper left quadrant with powerful pan and zoom capabilities and the ability to toggle between the native image with no analysis and the complete CytoSavvy analytics package.

The analytics row on the bottom of the dashboard features clickable interactive graphics tied to targeted elements of the disease state. For kidney cancer, these graphics summarize the number of targeted nuclei found by grade including raw counts and a percentage distribution. Clicking on any of the bars displays the top 100 targeted elements (graded nuclei in this case) in the Gallery window in the upper right column of the dashboard. Item details allows the user to click on any targeted element in the image, while the results box displays a pro forma system grade, allowing the clinician or researcher to assign an final grade, approve the slide, and forward those results to a standard report or an internal reporting system integrated with the dashboard.

The Gallery window is activated by clicking a column of the bar chart and displays the top one-hundred targeted elements (graded nuclei in this case). Items in the gallery are sorted by proprietary figure of merit based, in this case, sorted on the nuclear size and the presence of chromatin and nucleoli. By clicking on the item, the system automatically zooms in on the item in the larger image. At this point, the clinician or researcher can turn off the analytics to examine the element and either move on to another gallery item or assign a final grade and approve the image. By providing objective data and a user-friendly interface, the goal of the kidney dashboard is to allow the researcher or clinician to reach a decision within three clicks.

