Statement of Purpose

Sanyam Kapoor — *MS in Computer Science Applicant*
sanyam@nyu.edu — *Computer Science Department at NYU Courant*

During my sophomore year, I co-founded a video startup called *StoryXpress* motivated by a strong interest to improve video creation technology at scale. I was responsible for building the Video Rendering Engine from scratch on top of *OpenGL*, a low-level graphics API specification. This included creation of 1D and 2D shapes compounded with the animation engine to produce visually compelling videos. The challenge was rendering videos on-demand in real-time. With memory overflows and runtime errors occurring often, I started investigating more efficient algorithms and data structures. After much research, I built the in-memory caching system to hold reusable data and invented heuristics to improve runtime. In one particular case, using triangle strips to generate shapes instead of quadrilaterals gave huge efficiency gains as lesser vertices were now needed to generate a mesh. The engine peaked at 150 High Definition videos per hour per machine sans hardware acceleration with error rate down to 1 in 400. *StoryXpress* was recognized as *NASSCOM Emerge 50* and also awarded the *Best Software Product in Student Innovation*.

Among interactions with many customers from the industry, one strikingly suggested the use of visual context from various media sources to generate relevant videos. This idea has since kept me engaged and been the source of my motivation to conduct research in *Computer Vision*.

Henceforth, I decided to pursue research work under Dr. Vineeth Balasubramanian at IIT Hyderabad. One of my first problems was *Partial Face Recognition* which aimed to improve object recognition in natural environments where performance gets affected by extraneous factors like shadows, illumination variations, and occlusions. I investigated and implemented a method which modeled images as graphs and the spatial similarity analysis of the same allowed grouping different pixels together to detect the face.

I extended this research in face recognition towards a project titled *eDrishti, Engagement Level Detection in Videos*. This demanded finer facial feature extraction to classify engagement levels of a MOOC video viewer. My model predicted two out of every three instances correctly. I discovered that predictions were failing because of unaccounted head movements leading to sub-optimal facial features. This led me to investigate object recognition in videos via the project titled *Sports Recognition via Dense Trajectory Features*. I used methods like *Histogram of Oriented Flows* and *Motion Body Histogram* for the optical flow analysis and trained the model to predict the sport being played.

These experiences have exposed me to interesting methodologies and new challenges have kept me captivated. An MS in Computer Science at New York University is my logical next step to carry this momentum forward.

My choice in coursework has been a conscious decision to supplement these research goals with strong theoretical foundations, the most integral being *Computer Vision*, *Numerical Linear Algebra*, *Soft Computing*

and *Predictive Analytics & Knowledge Discovery*. Laboratory courses have complemented my theory. I was part of the team that automated the *Timetable Management System - QuickSlots* for the Computer Science Department at IIT Hyderabad. Since most scheduling problems are *NP-hard*, we modeled the problem as a graph and built a set of constraints to allow for practical usage including semi-automated conflict resolution. This successful execution boosted my confidence in the ability to apply theory to challenging problems.

Ideas from the courses *Operating Systems* and *Compiler Design* have profoundly influenced me as well. One of the most powerful ideas to arise from the *UNIX* philosophy is the arrangement of all objects as a file. Finding synergy with my work at *StoryXpress*, I realized that the same design pattern could be effectively used for building applications. Since then, I maintain a blog series about engineering strategies I?ve learned while building production-level services. In one instance, I critiqued a deployment pattern called *Microservices* and suggested scenarios when one should migrate to such a pattern. This exercise has helped me achieve breadth of ideas and evolve as a pragmatic engineer.

Amid the same exercise, I came across a paper titled *Mesos: A Platform for Fine-Grained Resource Sharing in the Data Center* which proposed primitives for task and resource isolation in data centers. Impressed by the idea, I eventually went on to become a valuable contributor to the *Mesos Framework Development SDK* where I added *container* support. This experience gave me the opportunity to widen my scope and learn about challenges in distributed systems. Since problems in *Computer Vision* require large computational resources, I am optimistic about *Mesos* as a platform to run distributed machine learning algorithms at scale.

These experiences of conducting motivated research and the ability to apply in practice put me in a unique position to learn and contribute ideas among other highly qualified members of the community. New York University provides me an excellent environment among seasoned researchers. In continuation of my previous work, I am excited about the research being done at the Vision Learning Lab. I am particularly interested in solving challenging problems in Object Recognition. The opportunity to learn from courses like *Computer Vision* by Rob Fergus and *Advanced Computer Vision* by Davi Geiger will provide me the necessary depth to contribute novel ideas towards the same.

After I graduate, I aim to continue my research in *Computer Vision* (objects and their relational representations) and culminate it as relevant products in governance and logistics. New York University provides me with just the right set of opportunities to achieve my vision.